

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2001, 13:26:37 ; Search time 19.49 seconds
(without alignments)
2102.715 Million cell updates/sec

Title: US-09-271-584A-2
Perfect score: 2755
Sequence: 1 MLDLSVSKLPSTSDHASV.....FVPVPGSPTRPNPDLPSKA 538
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues
Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2136	77.5	457	2 T01804	Na+/H+-exchanging
2	583.5	21.2	703	2 T26529	hypothetical prote
3	564	20.5	633	2 S69734	hypothetical prote
4	544	19.7	569	2 T37706	probable sodium/hy
5	531	19.3	629	2 T22848	hypothetical prote
6	475	17.2	832	2 A40205	Na+/H+-exchanging
7	467	17.0	831	2 B40204	Na+/H+-exchanging
8	455.5	16.5	813	2 A46748	Na+/H+-exchanging
9	454.5	16.5	809	2 A46747	Na+/H+-exchanging
10	443	16.1	815	2 S17487	Na+/H+-exchanging
11	437	15.9	816	2 S16328	Na+/H+-exchanging
12	436	15.8	818	2 A48858	Na+/H+-exchanging
13	434.5	15.8	698	2 A57644	Na+/H+-exchanging
14	431	15.6	759	2 A46188	CAMP-activated Na+
15	430.5	15.6	717	2 C40204	Na+/H+-exchanging
16	424	15.4	478	2 T18746	Na+/H+-exchanging
17	423.5	15.4	820	2 A40204	Na+/H+-exchanging
18	420	15.2	822	2 S30198	Na+/H+-exchanging
19	406	14.7	798	2 T23539	hypothetical prote
20	391.5	14.2	634	2 T33528	hypothetical prote
21	385.5	14.0	651	2 T31869	hypothetical prote
22	381	13.8	375	2 B40205	Na+/H+-exchanging
23	366.5	13.3	660	2 T28016	hypothetical prote
24	327.5	11.9	609	2 S30910	Na+/H+-exchanging
25	311.5	11.3	494	1 F69355	Na+/H+ antiporter
26	289	10.5	684	2 T16072	hypothetical prote
27	279.5	10.1	531	2 D96827	protein F20B17.4 [
28	278.5	10.1	527	1 S75063	Na+/H+-exchanging
29	254.5	9.2	575	2 D96585	hypothetical prote

ALIGNMENTS

RESULT 1

T01804
Na+/H+-exchanging protein 3 homolog A_TM021B04.4 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 22-Oct-1999
C:Accession: T01804
R:Dante, M.; Wamsley, P.; Gibson, A.
submitted to the EMBL Data Library, June 1997
A:Description: The sequence of A. thaliana TM021B04.
A:Reference number: 214440
A:Accession: T01804
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-457 <DAN>
A:Cross-references: EMBL:AF007271; NID:g2191181; PID:g2191184; GSPDB:GN00063; ATSP:A-
A:Experimental source: cultivar Columbia
C:Genetics:
A:Gene: ATSP:A_TM021B04.4
A:Map position: 5
A:Introns: 55/3; 95/2; 128/1; 148/1; 163/3; 240/3; 256/2; 286/3; 320/2; 400/3; 421/3

Query Match	77.5%	Score	2136;	DB	2;	Length	457;
Best Local Similarity	95.6%	Pred. No.	9.8e-155;				
Matches	430;	Conservative	0;	Mismatches	0;	Indels	20;
Gaps	2;						
QY	1	MLDSLVSKLPSTSDHASVVALNLFVALLCACIVLGHLEENRWNNESITALLIGLGTG	60				
Db	1	MLDSLVSKLPSTSDHASVVALNLFVALLCACIVLGHLEENRWNNESITALLIGLGTG	60				
QY	61	VVILLISKGKSHLLVSEDLFFYILLPPIINAGFQVKKQFFNFVTIMFGAVGTII	120				
Db	61	VVILLISKGKSHLLVSEDLFFYILLPPIINAGFQVKKQFFNFVTIMFGAVGTII	120				
QY	121	SCITISLGVTFQFKKLDIGTFLDGLDYLAIGAFAAADSVCITLQVNLQDETPLYSLVFE	180				
Db	121	SCITISLGVTFQFKKLDIGTFLDGLDYLAIGAFAAADSVCITLQVNLQDETPLYSLVFE	180				
QY	181	GVVNDATSVVFNAIQSFDTLHNEAAPHLLGNFLYLLSTLLGAA-----	228				
Db	181	GVVNDATSVVFNAIQSFDTLHNEAAPHLLGNFLYLLSTLLGAAVPSLSSLPFFL	240				
QY	229	TGLISAYIKKLYPG-----RHSTDREVALMMLMAYSLMELFDLSGLITVFFCG	280				
Db	241	TGLISAYIKKLYPGPHINCHRHSTDREVALMMLMAYSLMELFDLSGLITVFFCG	300				
QY	281	IVMSHYTHWNTSSRTTKHTFATLSFLAETFIYLVGMADALDKWRSVSDPTGTSIA	340				
Db	301	IVMSHYTHWNTSSRTTKHTFATLSFLAETFIYLVGMADALDKWRSVSDPTGTSIA	360				
QY	341	VSSILMGLVWVGRAAFVFPPLSFLSNLAKKNOSEKINFNMQVIVWSGLMRGVSNALAYN	400				
Db							

Db 361 VSSILMLGVYVGRFAAFVPLSFLSNLAKKQSEKINFNQVVMWGLMRGAVSMALAYN 420

QY 401 KETRAHGTDRGNAMITSTITVCLFSTVY 430

Db 421 KETRAHGTDRGNAMITSTITVCLFSTVY 450

RESULT 2

hypothetical protein Y18D10A.6 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T26529

R:Harri, B.

submitted to the EMBL Data Library, December 1998

A:Reference number: 220226

A:Accession: T26529

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-703 <WIL>

A:Cross-references: EMBL:AL034393; PIDN:CAA22320.1; CESP:Y18D10A.6

A:Experimental source: clone Y18D10A

C:Genetics:

A:Gene: CESP:Y18D10A.6

A:Introns: 23/3; 56/1; 103/2; 246/3; 352/3; 492/2; 578/3

Query Match 21.2%; Score 583.5; DB 2; Length 703;

Best Local Similarity 34.6%; Pred. No. 1.4e-36;

Matches 150; Conservative 79; Mismatches 160; Indels 45; Gaps 12;

QY 25 LFVALLCACIVLGHLEENRMNESITALLIGLGTGTITLISKSKSHLLVSEDLFFI 84

Db 138 LFVIMLATVVMHVLIVSKIHMPESLAIVAGALIG-STLSYSRDWSIEALSPDVFLL 196

QY 85 YLLPPIIFNAGFQVKKQFFRNFTYIMLFGAVGTITISLGTVOFFKLDIGFDLG 144

Db 197 VLLPPIIFNAGFQVKKQFFRNFTYIMLFGAVGTITISLGTVOFFKLDIGFDLG 256

QY 145 DYLAIGALFAATDSVCTLOVNODET-PLLYSLVFGGVDNATSVVY-----F 192

Db 257 ECFAPAAISADVDPGLTAFQAVKVESLYMLVFGESMLNDANSIVLAAVALAKHAKPSP 316

QY 193 NAIQSFOLTHLHNEAFAHLLGNFLYLLSTLGAATGLISAYIKKLYFORHSTDEVA 252

Db 317 NSLPASEII---TSAP---VTTEFFFSACLGVCIGLSSALLFKHVDL-RKTPSLEPA 368

QY 253 LMMLMAYLSYMLAEFLDLSGILTVFCGIVMSHYTHWNTSSRTTKHTFAILSLAET 312

Db 369 LLLIFSYIPYGAALDLSGIMAILFCGISMQFTRHNVPITQITFRHTFRTISFVAET 428

QY 313 FIFLYVGM-----ALDIDKWRVSVDTPGTSIAVSSILMGLVMVGRFAAFVPLSFLSLAK 368

Db 429 SFYFAYIGMAFFITILNFPW-----LIFWSV--LCILGRACNVFPLAYLVNQCR 476

QY 369 KNOSEKINFNQVVMWGLMRGAVSMALAYNKFTRAGHTDVRGNAMITSTITVCLFST 428

Db 477 KD--VOISMKQIIMWFSG-MRGAVCFALVLYM-----DLQEKKSILLITVFLILFTT 528

QY 429 VVFGMLTKPLISYL 442

Db 529 IFLGGSALPFIISFI 542

RESULT 3

S69734

hypothetical protein YDR456w - yeast (*Saccharomyces cerevisiae*)

C:Species: *Saccharomyces cerevisiae*

C:Date: 22-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 23-Mar-2001

C:Accession: S69734

R:Dieckhoff, F.S.

submitted to the EMBL Data Library, August 1995

A:Description: The sequence of *S. cerevisiae* lambda 3641 and cosmids 9461, 9831, and 9411

A:Reference number: S69555

A:Accession: S69734

A:Molecule type: DNA

A:Residues: 1-633 <DI>

A:Cross-references: EMBL:U33007; NID:9927685; PID:9927695; GSPDB:GN00004; MIPS:YDR456

C:Genetics:

A:Gene: MIPS:YDR456w

A:Map position: 4R

C:Superfamily: hypothetical protein yv9p

Query Match 20.5%; Score 564; DB 2; Length 633;

Best Local Similarity 26.8%; Pred. No. 3.8e-35;

Matches 164; Conservative 121; Mismatches 200; Indels 128; Gaps 21;

QY 14 TSDHASVVALNLFVALLCACIVLGHLEEN--RMNNEISITALLIGLGTGTITLISKSKS 71

Db 55 TEEMFSSWALFTMLLLISALWSSYYLTQKRIRAVHETVLSIFYGMVIGLIIIRMSPGHYI 114

QY 72 SHLVFESDLFFIYLLPPIIFNAGFQVKKQFFRNFTYIMLFGAVGTITISLGTVO 131

Db 115 QDVTVTNSSYFFNVLLPPIILNSGVYELNQVNFNNMLSILIFAIPTGTFISAVVIGI-ILY 173

QY 132 FRKKDICTFDL--GDYLAIGALFAATDSVCTLOVNODET-PLLYSLVFGGVDNAT 188

Db 174 IWTFLGESIDISFADAMSVGATLSATDPVITISFNAYKVDPKLYIIFGESLNDALS 233

QY 189 VVFNIAIOSFDLTHLHNEAFAHLLGNFLYLLSTLGAATGLISAYIKKLYFORHSTD 248

Db 234 IVNFETCORFHQOPATFSVFEAGLFLMTFESVLLIGVILVALLKHTHIRRY-PQ 292

QY 249 REVALMMLAYLSYMLAEFLDLSGILTVFCGIVMSHYTHWNTSSRTTKHTTATLSF 308

Db 293 IESCLILLIAYESYFSGCHMSGIVSLLECGITLKHAYYNNMSRSQITIKYIQLLAR 352

QY 309 LATFTFLYVGMADLIDKWRVSVDTPGTS-----TAVSSILMGLVMVGRFAAFVFP 359

Db 353 LSENFIFYLGLLELF-----TEVELYKPLLIIVAAI---SICVARNCAVEFP 396

QY 360 LS-----FLSNLAKKNS--EKINFNMQVVMWGLMRGAVSMALAYN---- 400

Db 397 LSQFVWNIYRVKTIKNSGITGENISVDEIPYNYQMTFAGL-RGAVGVALAIGIQE 455

QY 401 -KTRAGHTDVRGNAMITSTITVCLFSTVVFVGMGLTKPLISYLLPHONATSMLEDDNT- 458

Db 456 YKFT-----LLATVIVVVLTVIIFGTTAGMLEVL---NIKTGCSIEDTS 499

QY 459 -----PKSIHI-----PLLDQDSFIEPSGHNHVPDPDSIRGELT 492

Db 500 DDEFDIEAPRAINLLNGSSIoTDLGPYSDNNSPDISIDQF-AVSSNKNLPNNISTTNGNT 558

QY 493 -----RPTRTV-----HYWRFQFDDSPMRPVFGGRGVPFV 523

Db 559 FGGLNETENTSPNPARSSMDKRNLRDLKLTFTNSDSQWQNFQDEQVLKPVFLD-NVSPSL 617

QY 524 PGSPTRNPPDLS 536

Db 618 QDSATQ-SPADES 629

RESULT 4

T37706

probable sodium/hydrogen exchanger - fission yeast (*Schizosaccharomyces pombe*)

C:Species: *Schizosaccharomyces pombe*

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000

C:Accession: T37706

R:Murphy, L.; Harris, D.; Wood, V.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, June 1997

A:Reference number: Z21738

A:Accession: T37706

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-569 <MUR>

```
Query Match      19.3%; Score 531; DB 2; Length 629;
Best Local Similarity 32.2%; Pred. No. 1.2e-32;
Matches 136; Conservative 91; Mismatches 129; Indels 66; Gaps 14;

QY       76 VFSEDLFFYLPLPIFNAGFQVKKKGFERNFTVIMLFAGVTIISC-----TI 124
          |||| :||| :|||||:: :||| :||| :||| :||| :||| :|||
```

RESULT 9		A:Status: preliminary; translated from GB/EMBL/DBDJ	
A46747		A:Molecule type: mRNA	
Na+/H+-exchanging protein NHE-2 - rabbit		A:Residues: 1-815 <FLI>	
C:Species: Oryctolagus cuniculus (domestic rabbit)		A:Cross-references: GB:S68616; NID:g544775; PIDN:AAC60606.1; PID:g544776	
C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 28-Feb-1997		R:Sardet, C.; Franchi, A.; Pouyssegur, J.	
C:Accession: A46747		Cell 56, 271-280, 1989	
R:Tse, C.M.; Levine, S.A.; Yun, C.H.; Montrose, M.H.; Little, P.J.; Pouyssegur, J.; Dono		A:Title: Molecular cloning, primary structure, and expression of the human growth fac	
J. Biol. Chem. 268, 11917-11924, 1993		A:Reference number: A31311; MUID:89106219	
A:Title: Cloning and expression of a rabbit cDNA encoding a serum-activated ethylisopropr		A:Accession: A31311	
A:Reference number: A46747; MUID:93280159		A:Molecule type: mRNA	
A:Accession: A46747		A:Residues: 1-814 /SNARASQRSLPHRLFHQRCWGLPFLTRIGPAPPPTAWQLGPPQPHQSSSPASREASSPP	
A:Status: preliminary		C:Genetics:	
A:Molecule type: nucleic acid		A:Gene: GDB:SLC9A1; APNH; NHE1	
A:Residues: 1-809 <TSE>		A:Cross-references: GDB:119683; OMIM:107310	
A:Experimental source: ileal villus cells		A:Map position: lp36.1-lp35	
A>Note: sequence extracted from NCBI backbone (NCBIN:133350, NCBIP:133351)		C:Keywords: glycoprotein; transmembrane protein	
		F:75,370,410/Binding site: carbohydrate (Asn) (covalent) #status predicted	
		Query Match 16.1%; Score 443; DB 2; Length 815;	
		Best Local Similarity 26.0%; Pred. No. 8.1e-26;	
		Matches 151; Conservative 99; Mismatches 205; Indels 126; Gaps 22;	
QY	7 SKPLSLSDHASV---VALNLFVALLCACIVLGHLLLEE-NRWNNESITALLIGLGTGVT 62	QY	23 LNLFVALLCACIVLGHLLLEE-NRWNNESITALLIGLGTGVTILLISKSSHLVFSDEL 81
Db	66 SRLP-VFTLDYPHVQIPFEITLILLASLAKIGFHLVHKLPITVPESCLLIWVGLLGGI 124	Db	105 ISLWILLACLAKMKIGFHVPTISSIVPESCLLIWVGLLVGGI---KGVETPPPLQSDV 160
QY	63 ILLISKSSHLVFSDELFIYLLPPIIFNAGQVKKKQFFNFVIMLFGAVGTIIC 122	QY	82 FIYLLPPIIFNAGQVKKKQFFNFVIMLFGAVGTI-----ISCTIISLGVTQFF 133
Db	125 IFGVDE-KSPAM--KTDVDFLLPPIVDLADGYFMPTRPFENLGTIFWYAVVGTLWS 181	Db	161 FFLFLLPPIILDAGYFLPLRQFTENLGTILFVAVVGTLWNAFFLGLGVAVCLVGGQ-- 218
QY	123 TIISLGVTOFFKLDIGTDLGY----LAIGAIFAATDSVCTLOVL-NODETPLLYSL 176	QY	134 KKLDIGTDLGYLAIGAIFAATDSVCTLOVLNODE-TPLLYSLVFGGVNDATSVVVF 192
Db	182 --IGIGVS-LFGICQIEAGLSQITLQNLFLGSLISAVDPVAVLAVFENIHVNEQLYL 238	Db	219 ----INNIGLNLFLGSLISAVDPVAVLAVFEEIHELHILVFGESLLNDVTVVLY 274
QY	177 VFGGVNDATSVVVFNAIQSF-DLTHLNHEAAFLHNLGFLYLLSTLIGAAATGLISAY 235	QY	193 NAIQSFDLTHLNHEAAFLHNLGFLYLLS---TLGGAATGLISAYVIKKLYFGRHSTD 249
Db	239 VFGESLNDATVVLYNLKFSFCOMKTIEDIVFAGIANEFVVGIGVIGLIGFLGFIAP 298	Db	275 HLFEF--ANYEHVGVIVDIFLGLFSLFVVALGVGVVGVVIAFTSR---FTSHIRVI 329
QY	236 VIKKLYFGRHSTDREVALMMLAYSLYMLAELFDLSGILTVFPCGIVMSHYTHNVTESS 295	QY	250 EVALMMLAYSLYMLAELFDLSGILTVFPCGIVMSHYTHNVTESSRITTKHTFATLSFL 309
Db	299 TTR---ETHNIRVIEPLFVLYSLYITAEFHLSCIMAITACAMTKNRYEENVSOKS 355	Db	330 EPLFVLYSLYMAFLSAELFHLGIMALIASGVVMPYVEANISHKSHTTIKYFLKMNSV 389
QY	296 RITTKHFTATLSFLAEFIPLYVGMALDID---KRSVSDTPGTSIATVSSILLMGLVMVG 352	QY	310 ATEFIPLYVGMALDIDKRSVSDTPGT-----STAVSSILLMGLVMGRAAFVPLSLS 364
Db	356 YTIKYPKMLSSVSETLIFIFMGVSTVGKNHEWNAFVCT-----LAFCLIW 404	Db	390 SETLIFIFLG-----VSTVAGSHHNNWTVISTLL--FCLIARVILGVLGTWFI 436
QY	353 RAAPVPLSFLSLNLAQKQSEKINFNMQVVIWMSGLMRGAVSMALAY----NKFTTRAGHT 408	QY	365 NLAKNQSEKINFNMQVVIWMSGLMRGAVSMALAY----NKFTTRAGHTDVRGNAIMITST 420
Db	405 RALGVFVLRVINWFR---DIPLTFKQDFIIAYGGL-RGAICFALVFLPAAVFPF--- 456	Db	437 N---KFRIVKLTPKQDFIIAYGGL-RGAIAFSLGYLLDKKHPPMCD-----LFLTAI 484
QY	409 DVRGNAIMITSTVCLFSTVFGMLTKPLISYL-LPHQNATTSMLSDNTPKSIHIPL 467	QY	421 ITVCLFSTVFGMLTKPLISYLLPHQNATTSMLSDNTPKSIHIPLLDODSFIEPSSGHNH 480
Db	457 ----KKLFIATAIIVIFFFVILGITIRPLVFEFLDKRSNKKQQAQVSEE-----IHCRFF 507	Db	485 ITVIEFTVFGQMTIRPLVDLAVKKQKQETKR---SINEEHTQFLD-----HL 530
QY	468 DQ-DSFTEPSGNINVRPDSIRGFLTRPTVHYWYR---QFDDSEFMR 511	QY	481 VPRPDSIRGFLTRPTVHYWYR---QFDDSEFMRPVF----- 514
Db	508 DHVKTGIEDVCGH-----WGHNFRDKEFKFDDKIYLR 539	Db	531 LTGIEDICGHY-----HHHWKDKLNRFNKYVKKKLIAGERSKEPOLIAPYHKMEMKQ 584
RESULT 10		QY 515 -----GGRGFVPFV-----PGS-PTERNPPDLSK 537	
I57487		Db 585 ATELVESGGMGKIPSAVSTVSMQNIHPKSLSPERITLPALSK 625	
Na+/H+-exchanging protein, amiloride-sensitive - human		RESULT 11	
N:Alternate names: Na+/H+ antiporter; NHE-1		S16328	
C:Species: Homo sapiens (man)		Na+/H+-exchanging protein - rabbit	
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 22-Oct-1999		N:Alternate names: Na+/H+ antiporter; Na+/H+ exchanger; pH regulatory protein	
C:Accession: I57487; A31311		C:Species: Oryctolagus cuniculus (domestic rabbit)	
R:Fliegel, L.; Dyck, J.R.; Wang, H.; Fong, C.; Haworth, R.S.		C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 01-Dec-2000	
Mol. Cell. Biochem. 125, 137-143, 1993		C:Accession: S16328; S13926; S30602	
A:Title: Cloning and analysis of the human myocardial Na+/H+ exchanger.		R:Tse, C.M.; Ma, A.I.; Yang, V.W.; Watson, A.J.M.; Levine, S.; Montrose, M.H.; Potter	
A:Reference number: I57487; MUID:94111706		EMBO J. 10, 1957-1967, 1991	
A:Accession: I57487			

A;Species: picramandi
A;Molecule type: MRNA
A;Residues: 1-717 <ORL>
A;Cross-references: GB:M853301; NID:q205708; PID:AAA41703.1; PID:q205709

C:Keywords: transmembrane protein

```
Query Match      15.6%; Score 430.5; DB 2; Length 717;
Best Local Similarity 26.2%; Pred. No. 6.1e-25;
Matches 137; Conservative 102; Mismatches 189; Indels 95; Gaps 21;

QY 36 LCHLLLENRWNESTIALLIGTGTVTILLISKGSSHLVPSDDLFFIYLLPPIFNAG 95
Db 93 LPHL-----MPESCLLIIVGALVG-SIIFGTHKSP--VMDSSYFLYLLPPIVLES 143
QY 96 FOVKKKOFFRNFVIMLFGAVGTIISCTIISLGVTQFF----KKLDIGTFDLGDIYLAIGA 151
Db 144 YFMTPRPFFENIGSILWAGLALINA--FGIGLSLYFICQAKAFGLGIDINLLQNLLFGS 201
QY 152 IFAATDSVCTIQVLNQDE-TPLLYSLVFGEGVVDNATSVVFNAIQSFDLTHL--NHERA 208
Db 202 LISAVDPVAVLAVFEERARVNEQLYMMIFGEALLNDGISVLYNILLIAFTKHKFEDIEAV 261
QY 209 FHLIG--NFLYFLSTLLGAATGLISAVIKKLYFGRHSTDRVALMMLMAYLSYNLAE 266
Db 262 DILAGCARFVIVGGGVFFGIIIEGFIISAFITR---FTQNIISAIEPLIVFMFSYLSVIAAE 318
QY 267 LFDLSGILTVFCCGIVMSHYTNHNVTESSRIITKHTFATLSFLAETFIPLYVGMALDID 326
Db 319 TLYLSGILAITACAVTMKKYVEENVSTSYTTIKYFMKMLSSVSETLIFIFMGVSTVGKN 378
QY 327 ---KWRSDTPGTSIAVSSILMGLVMVCGRAAFVFLSFLSNLAKKNOSEKINFN--QV 381
Db 379 HEWNMAFVCF-----LAFQIWRKISVFTLFYS-----NQRTFFPSIKDQL 422
QY 382 VLVWSGLMRGAVSMALAY-----NKFTAGHTDVRGNAMITITVCLFSTVVFGLMTRP 437
Db 423 IIFYSG-VRGAGSFLAFLPLTLFPR-----KKLFVTATLVVTVYTFVFGITIGP 473
QY 438 LISYLLPHONATTSMLSDDNTPKSIHPLLDQDSFIEPSGMHNVPRDPSIRGELTRPTT 497
Db 474 LVRYLDVRKTNKESINEE-----LHRLMD-----HLKAGIEDYCG-----QW 512
QY 498 VHYH----WROFDDSFMRPVFGGGEVPEVPGSPTERNPPDLS 536
Db 513 SHYQVRDKFKFDHRYLRKIL-----IRRNQPKSS 542
```

Search completed: November 17, 2001, 13:28:25
Job time: 108 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2001, 13:27:12 ; Search time 13.15 seconds
(without alignments)
1401.479 Million cell updates/sec

Title: US-09-271-584A-2
Perfect score: 2755
Sequence: 1 MDSLVKPLSLSTDSHVS.....FVPFVGSPTERNPPDL SKA 538

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	581	21.1	669	1 NAH6_HUMAN	Q92581 homo sapien
2	564	20.5	633	1 NAH2_YEAST	Q04121 saccharomyc
3	475	17.2	832	1 NAH3_RABIT	P26432 oryctolagus
4	467	17.0	831	1 NAH3_RAT	P26433 rattus norv
5	466.5	16.9	896	1 NAH5_HUMAN	Q14940 homo sapien
6	466.5	16.9	898	1 NAH5_RAT	Q920x2 rattus norv
7	462	16.8	834	1 NAH3_HUMAN	P48764 homo sapien
8	459	16.7	839	1 NAH3_DIDMA	Q28362 didelphis m
9	455.5	16.5	813	1 NAH2_RAT	P48763 rattus norv
10	454.5	16.5	809	1 NAH2_RABIT	Q28482 oryctolagus
11	452.5	16.4	812	1 NAH2_HUMAN	Q9ubyo0 homo sapien
12	443	16.1	815	1 NAH1_HUMAN	P19634 homo sapien
13	437	15.9	816	1 NAH1_RABIT	P23791 oryctolagus
14	436	15.8	818	1 NAH1_PIG	P48762 sus scrofa
15	431	15.6	759	1 NAH6_ONCMY	Q01345 oncorhynch
16	431	15.6	820	1 NAH1_RAT	P26431 rattus norv
17	430.5	15.6	717	1 NAH4_RAT	P26434 rattus norv
18	429	15.6	817	1 NAH1_BOVIN	Q28036 bos taurus
19	424	15.4	820	1 NAH1_MOUSE	Q61165 mus musculu
20	420	15.2	822	1 NAH1_CRIGR	P48761 cricetus
21	327.5	11.9	609	1 NAH1_CABEL	P35449 caenorhabdi
22	202	7.3	549	1 YJCE_ECOLI	P32703 escherichia
23	190	6.9	542	1 YMB7_MYCTU	Q50678 mycobacteri
24	160.5	5.8	578	1 YCGO_ECOLI	P76007 escherichia
25	145	5.3	383	1 YAGO_ENTHR	P26235 enterococcu
26	144.5	5.2	426	1 Y057_METJA	Q60362 methanococc
27	140.5	5.1	759	1 NAH2_SCHPO	Q14123 schizosacch
28	128.5	4.7	422	1 YF21_METJA	Q58916 methanococc
29	121	4.4	407	1 GLUP_HELPY	Q25788 helicobacte
30	120.5	4.4	808	1 NAH2_ZYGRO	Q42701 zygosacchar
31	119.5	4.3	581	1 ATKA_ANASL	Q9f6x2 anasana sp
32	117.5	4.3	430	1 RFBX_SALTY	P26400 salmonella
33	117	4.2	558	1 YBAL_ECOLI	P39830 escherichia

34	116.5	4.2	614	1 NUOL_BUCAI	P57262 buchnera ap
35	113.5	4.1	549	1 COX1_LEITA	P14544 leishmania
36	113.5	4.1	549	1 COX1_TRYBB	P04371 trypanosoma
37	113	4.1	407	1 GLUP_HELPY	Q92k41 helicobacte
38	113	4.1	519	1 NU4M_FODAN	P15582 podospora a
39	111.5	4.0	459	1 NU4M_RABIT	O79436 oryctolagus
40	111.5	4.0	468	1 NAH_SCHPO	P36606 schizosacch
41	111	4.0	553	1 NUOM_MYCTU	O53307 mycobacteri
42	111	4.0	815	1 AOX1_AERPE	Q9ydx6 aeropyrum p
43	110.5	4.0	459	1 NU4M_HIPAM	O92zy2 hippopotamu
44	107.5	3.9	402	1 Y718_METJA	O58128 methanococc
45	107.5	3.9	514	1 COX1_PELSU	O79672 pelomedusa

ALIGNMENTS

RESULT 1
NAH6_HUMAN
ID NAH6_HUMAN STANDARD; PRT; 669 AA.
AC Q92581;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE SODIUM/HYDROGEN EXCHANGER 6 (NA(+)/H(+) EXCHANGER 6) (NHE-6).
GN SLC9A6 OR NHE6 OR KIAA0267.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98175963; PubMed=9507001;
RA Numata M., Petrecca K., Lake N., Orłowski J.;
RT "Identification of a mitochondrial Na⁺/H⁺ exchanger.";
RL J. Biol. Chem. 273:6951-6959(1998).
RN [2]
RP SEQUENCE OF 4-669 FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=97191544; PubMed=9039502;
RA Nagase T., Seki N., Ishikawa K.-I., Ohira M., Kawarabayashi Y.,
Chara O., Tanaka A., Kotani H., Miyajima N., Nomura N.;
RT "Prediction of the coding sequences of unidentified human genes. VI.
The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by
analysis of cDNA clones from cell line KG-1 and brain.";
RL DNA Res. 3:321-329(1996).
CC -I- FUNCTION: ELECTRONEUTRAL EXCHANGE OF PROTONS FOR NA⁺ AND K⁺ ACROSS
THE MITOCHONDRIAL INNER MEMBRANE. CONTRIBUTES TO ORGANELLAR VOLUME
AND CALCIUM HOMEOSTASIS.
CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRION.
CC -I- TISSUE SPECIFICITY: UBIQUITOUS; BUT IS MOST ABUNDANT IN
MITOCHONDRIUM-RICH TISSUES SUCH AS BRAIN, SKELETAL MUSCLE, AND
HEART.
CC -I- SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF030409; AAC39643.1; -;
DR EMBL; D87743; BAAI3449.1; -;
DR MTM; 300231; -;
DR InterPro; IPR000676; -;
DR InterPro; IPR002090; -;
DR Pfam; PF00999; Na.H.Exchanger; 1.
DR PRINTS; PR01084; NAHEXCHNGR.
DR PRINTS; PR01086; NAHEXCHNGR6.
DR Transmembrane; Sodium transport; Transport; Symport; Mitochondrion.
KW TRANSMEM 28 48 POTENTIAL.
FT

FT	TRANSMEM	71	91	POTENTIAL.	
FT	TRANSMEM	103	123	POTENTIAL.	
FT	TRANSMEM	142	162	POTENTIAL.	
FT	TRANSMEM	179	199	POTENTIAL.	
FT	TRANSMEM	220	240	POTENTIAL.	
FT	TRANSMEM	246	266	POTENTIAL.	
FT	TRANSMEM	292	312	POTENTIAL.	
FT	TRANSMEM	340	360	POTENTIAL.	
FT	TRANSMEM	382	402	POTENTIAL.	
FT	TRANSMEM	404	424	POTENTIAL.	
FT	TRANSMEM	447	467	POTENTIAL.	
FT	TRANSMEM	483	503	POTENTIAL.	
SQ	SEQUENCE	669 AA;	74161 MW;	F6416596229F2639	CRC64;
Query Match 21.1%; Score 581; DB 1; Length 669;					
Best Local Similarity 30.0%; Pred. No. 6.9e-32;					
Matches 166; Conservative 102; Mismatches 173; Indels 112; Gaps 20;					
QY	23	LNLFVALLCACIVLGHLEEN--RWNESITALLIGLGTGVITLL-----	65		
Db	74	LLIFILLTUTLITWLKHRRARFHEHGLMIYGLLVGLVLRVYGHVPSDVNNVTLSC	133		
QY	66	-ISKGSHLLVSEDLFFYLLPPIIFNAGFQVKKQFFRNFVTIMLFGAVGTIISCTI	124		
Db	134	EVQSSPTLLVTFDPEVFVFNILLPPIIFVAGYSLKRRHFFERNLGSILAYFLGTAISCFV	193		
QY	125	IS---LGVTOFFKKLD--IGTFDLGDYLAIGAIFAAATDSVCYLQVLNQDETPL-LYSLVF	178		
Db	194	IGSIMYCVTLMLKVTGQAGDFYFTDCLLFGAIVSATDPVTYLAIFHELOVDVELYALLF	253		
QY	179	GEGVNDATSVVVFNAI-----QSFDLTHLNEAAFLHLLGNFLYLFLLSTLLGAA	228		
Db	254	GESLVDAVAIVLSSSVAVIQAGDNSHTFDVT-----AMFKSIGIFLGIFSGSPANGAA	308		
QY	229	TGLISAYVIK--KLYFGRHSTDEVALMMLMAYLSVYMLAEFLDLSGLIVTFPGGIVMSHY	286		
Db	309	TGVVTVLVTFTKL---REFQLLETGLFLLMSWSTFLAEAWGTGVVAVLFCGIIQAHY	365		
QY	287	TWHNVTESSRTTKHTATLSFLAEAFIFLYVGDMDALDIDKWSRSDTPTGTSVAVSILM	346		
Db	366	TYNLSSTOHRQTKFELLNLAENFIPIFISGLTFL-----TFGNHVFNTFFVY	415		
QY	347	G---LVMWGRAAFVPLSFLSNLAKKNOSEKINFNMQVLIWWSGLMRGAVSMALAYNKFT	403		
Db	416	GAFVAIFLGRAANIYPLSLLNLGRS---KIGSNFOHMMFAGL-RGAMAFALAIR---	468		
QY	404	RAGHTDVRGNAMITSTITVCLFSTVYVFGMLTKPLISYLLPHQNATTSMLSDONTPKSIH	463		
Db	469	---DTATYAROMFSTLLIVFTVWVFG-----GGTAMLS-----CLH	505		
QY	464	lPL---LDODSFIEPSGNHNVPDPDSIGFTRPTRTVHYVYRQFDDSFMRPVFGGRGFV	520		
Db	506	IRVGDSQOEHLGVPENERRTTKAES--AWLFR-----MWYFNHNLKPLL-----	550		
QY	521	PFVPGSPTEPNPP	533		
Db	551	-----THSGPP	556		
RESULT 2					
ID	NAH2_YEAST	STANDARD;	PRT;	633 AA.	
AC	Q04121;				
DT	30-MAY-2000 (Rel. 39, Created)				
DT	30-MAY-2000 (Rel. 39, Last sequence update)				
DT	30-MAY-2000 (Rel. 39, Last annotation update)				
DE	MITOCHONDRIAL SODIUM/HYDROGEN EXCHANGER (MITOCHONDRIAL NA(+)/H(+)				
DE	EXCHANGER).				
GN	NAH2 OR NHX1 OR YDR456W OR D9461.40.				
OS	Saccharomyces cerevisiae (Baker's yeast).				
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;				
OC	Saccharomycetales; Saccharomycetaceae; Saccharomyces.				

```

Db 293 IESCLILLIAYESVFFSGCHMSGIVSLFLFCGITLKHAYAYNMSRRSQITIKYIFOLLAR 352
Qy 309 LAETFFIYVGMADALDIDKWSRSDTPGTS-----IAYSSILMLGMVWGAAFAVFP 359
Db 353 LSENFIIYGLLELF-----TEVELVYKPLLIIVAAI---SICVARNCAVFP 396
Qy 360 LS-----FLSNLAKNOS--EKINFNMQVVIWMSGLMRGAYSMALAYN----400
Db 397 LSQFVNMIYRVKTRISMSGITGENISVPDEIPYQMTFTWAGL-RGAVGVALALGIQGE 455
Qy 401 -KTRAGHTDVRGNAMITITIVCLFSTVVFGLTKPLISYLLPHONATTSMLSDNT- 458
Db 456 YKET-----LLATVLVVVTLVIIFGTTAGMLEVL----NKTCGISEEDTS 499
Qy 459 -----PKSIHT-----PLLDQDSFIEPSGNHNPVPRDSIRGFILT 492
Db 500 DDEFIEAPRAINLLNGSSIQDTGLGYPDSNNSPDISIDQF-AVSSKNKLPNNISTTGGNT 558
Qy 493 -----RPRTRV-----HYWRQFDDSPMRPVFGGRGFPVPV 523
Db 559 FGGLNETENTSPNPARSMDKRNRLKGLTIFNSDSQWFEQVLKPVFLD-NVPSPL 617
Qy 524 PGSPTEPNPDL 536
Db 618 QDSATQ-SPADFS 629

RESULT 3
ID NAH3_RABIT STANDARD; PRT; 832 AA.
AC R26432;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE SODIUM/HYDROGEN EXCHANGER 3 (NA(+)/H(+) EXCHANGER 3) (NHE-3).
GN SIC9A3 OR NHE3.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=NEW ZEALAND WHITE; TISSUE=ileal villus, and kidney cortex;
RX MEDLINE=92250340; PubMed=1374392;
RA Tse C.-M., Brant S.R., Walker S.S., Pouyssegur J., Donowitz M.;
RT "Cloning and sequencing of a rabbit cDNA encoding an intestinal and
RT kidney-specific Na+/H+ exchanger isoform (NHE-3).";
RL J. Biol. Chem. 267:9340-9346(1992).
CC -!- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED
CC BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL
CC CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD
CC SODIUM ION CHEMICAL GRADIENT. PLAYS AN IMPORTANT ROLE IN SIGNAL
CC TRANSDUCTION.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: INTESTINAL AND KIDNEY SPECIFIC. MOST ABUNDANT
CC IN KIDNEY CORTEX, FOLLOWED EQUALLY BY ILEUM AND ASCENDING COLON,
CC THEN KIDNEY MEDULLA AND JEJUNUM. IS ABSENT FROM DUODENUM AND
CC DESCENDING COLON.
CC -!- PTM: PHOSPHORYLATED (POSSIBLE).
CC -!- SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.
CC -!- CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC
CC DOMAINS IN THE NA(+)/H(+) EXCHANGERS VARY AMONG AUTHORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M87007; AAA31420.1; -
DR PIR; A40205; A40205.

```

```

DR InterPro: IPR000676; -
DR InterPro: IPR001118; -
DR Pfam: PF00999; Ne_H_Exchange; 1.
DR PRINTS; PR01084; NAHEXCHNGR.
DR PRINTS; PR01087; NAHEXCHNGR3.
KW Transmembrane; Glycoprotein; Sodium transport; Transport; Symport;
KW Multigene family; Phosphorylation; Polymorphism.
FT DOMAIN 1 11
FT TRANSMEM 12 27
FT DOMAIN 28 59
FT TRANSMEM 60 79
FT DOMAIN 80 81
FT TRANSMEM 82 101
FT DOMAIN 102 110
FT TRANSMEM 111 130
FT DOMAIN 131 134
FT TRANSMEM 135 154
FT DOMAIN 155 180
FT TRANSMEM 181 200
FT DOMAIN 201 209
FT TRANSMEM 210 229
FT DOMAIN 230 249
FT TRANSMEM 250 269
FT DOMAIN 270 298
FT TRANSMEM 299 319
FT DOMAIN 320 339
FT TRANSMEM 340 359
FT DOMAIN 360 366
FT TRANSMEM 367 385
FT DOMAIN 386 435
FT TRANSMEM 436 455
FT DOMAIN 456 832
FT CARBOHYD 325 325
FT VARIANT 144 144
SQ SEQUENCE 832 AA; 92748 MW; 8C8BB7C296CF8740 CRC64;

```

```

Query Match 17.2%; Score 475; DB 1; Length 832;
Best Local Similarity 29.18; Pred. No. 1.le-24;
Matches 148; Conservative 97; Mismatches 196; Indels 68; Gaps 19;

Qy 20 VVALNLFVALLCACIV-LGHLEENRMWNESITALLGLGTGVITILLISKSKSHLLVF- 77
Db 56 IIALWLVASLAKIVFHLSH--KVTSVVPESALLVLGLVGIVL-----AADHIAST 108
Qy 78 -SEDLFFIYLLPPIIFNAGFOVKKKQFPRVFTMLFGAVGTIISCTIISLGVTQFPFKL 136
Db 109 LTPTVFFYLLPPIVDAGYFMPNRLFESNLGSILLVAVGVTVWNAATTGLSLYGVFLSG 168
Qy 137 DIGTFDLG--DYLAIGAFAATDSVCTLOVUNQ--DETPLLYSLVFGGVNDATSVVVEN 193
Db 169 IMGELKGLDLDFLFGSLIAADVPVAVLAVFEVHVNEVLIIVFGESLLNDVAVTVLYN 228
Qy 194 AIQSFDTLHNHEAFAHLLGNFLYLFLSTLLGAATGLISAYVKKL-YFGRHSTDEVA 252
Db 229 VFQSVTLGGDKVTGDCVKGVISFVVVS-LGGTGVGVFAFLLSLVTRTKHVRVIEPG 287
Qy 253 LMMLMAYLSYMLAEFLDLSGILTVFFCGIVMSHTVHWNTVESSRIITTKHTFATLSFAET 312
Db 288 FVFIISLSYLTSEMLSSLSILAITFCGICQKYVKANISEQSATTVRYTMKMLASGET 347
Qy 313 FIFLYVGMDALIDKWSVSDTPCTSTAVSSILMGLVMVGRAAFVFPFLSLSLAKKNOS 372
Db 348 IIFMFLGISAVDPLIW-----TWTAFVLTLL--FVSVFAIGVVLQTLNRYRMVOL 400
Qy 373 EKINFNMQVVIWMSGLMRGAVSMALAY-----NKFTRAGHTDVRGNAMITSTTVCFLST 428
Db 401 ELID---QVVMVSGGL-RGAVAFALVALLDGNK-----VREKNLFVSTTIIVFFTV 448
Qy 429 VVFGMLTKPLISYLLPHONATTSMLSDNTPK----STHIFLLDDQDSFIEPSGNHNPVPRD 485
Db 449 IFQGLTIKPLVQWLKVKR-----SEHREPFLKNEKLHGRAFD-----HILSAIE 491

```


QY 486 SIRGLTRPTRTVHY-----WRQFDDSPM 510
 ID NAH3 RAT STANDARD; PRT; 831 AA.
 AC P26433;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE SODIUM/HYDROGEN EXCHANGER 3 (NA(+)/H(+) EXCHANGER 3) (NHE-3).
 GN SLC9A3 OR NHE3.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Kidney;
 RX MEDLINE=92250539; PubMed=1577762;
 RA Orlowski J., Kandasamy R.A., Shull G.E.;
 RT "Molecular cloning of putative members of the Na/H exchanger gene
 family. cDNA cloning, deduced amino acid sequence, and mRNA tissue
 expression of the rat Na/H exchanger NHE-1 and two structurally
 related proteins";
 RT J. Biol. Chem. 267:9331-9339(1992).
 RL
 CC -!- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED
 CC BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL
 CC CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD
 CC SODIUM ION CHEMICAL GRADIENT. PLAYS AN IMPORTANT ROLE IN SIGNAL
 CC TRANSDUCTION.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -!- TISSUE SPECIFICITY: MOST ABUNDANT IN COLON AND SMALL INTESTINE,
 CC FOLLOWED BY KIDNEY AND STOMACH.
 CC -!- PTM: PHOSPHORYLATED (POSSIBLE).
 CC -!- SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.
 CC -!- CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC
 CC DOMAINS IN THE NA(+)/H(+) EXCHANGERS VARY AMONG AUTHORS
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M85300; AAA41702.1; -;
 CC PIR; B40204; B40204.
 CC InterPro: IPR000676; -;
 CC InterPro: IPR001118; -;
 CC Pfam: PF00999; Na_H_Exchange; 1.
 CC PRINTS; PR01084; NAHEXCHNGR.
 CC PRINTS; PR01087; NAHEXCHNGR3.
 CC Transmembrane; Glycoprotein; Sodium transport; Transport; Symport;
 CC Multigene family; Phosphorylation.
 KW
 FT DOMAIN 1 11 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 12 25 A (M1) HYDROPHOBIC.
 FT DOMAIN 26 51 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 52 70 B (M2) HYDROPHOBIC.
 FT DOMAIN 71 76 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 77 96 C (M3) (POTENTIAL).
 FT DOMAIN 97 109 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 110 130 D (M4) (POTENTIAL).
 FT DOMAIN 131 136 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 137 157 E (M5) (POTENTIAL).
 FT DOMAIN 158 177 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 178 199 F (M5A) (POTENTIAL).
 FT DOMAIN 200 207 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 208 229 G (M5B) (POTENTIAL).
 FT DOMAIN 230 249 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 250 271 H (M6) (POTENTIAL).
 FT DOMAIN 272 287 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 288 306 I (M7) (POTENTIAL).
 FT DOMAIN 307 337 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 338 359 J (M8) (POTENTIAL).
 FT DOMAIN 360 366 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 367 387 K (M9) (POTENTIAL).
 FT DOMAIN 388 402 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 403 423 L. HYDROPHOBIC.
 FT DOMAIN 424 432 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 433 453 M (M10) (POTENTIAL).
 FT DOMAIN 454 831 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 323 323 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 831 AA; 93105 MW; 77A4BF10DF99E3F CRC64;
 Query Match 17.0%; Score 467; DB 1; Length 831;
 Best Local Similarity 28.3%; Pred. No. 3.8e-24;
 Matches 144; Conservative 99; Mismatches 205; Indels 60; Gaps 18;
 QY 20 VVALNLFVALLCACIV-LGHLLLENRMNNEITALLIGLGTGVITILLISKRSKSHLLVF- 77
 Db 54 IIALWILVASLAKIVFHLSH--KVTSVVPESALLIVGLVGLGIV-----WAADHIAST 106
 QY 78 -SEDLFIYLLPPIFNAGQVKKQFFRNFTVIMLFGAVGTIISCTIISLGVTFQFKKL 136
 Db 107 LPTLTFYLLPPIVLDAGYFMPNRLFFGNLTGILLYAVIGTITWNAATTGLSLGVFLSG 166
 QY 137 DIGTFDLG--DYLAIGAIFAATDSVCTQLVNLQ-DETPLLYSLVFGEVGVNDATSVVVEN 193
 Db 167 IMGELKTGLDLDLFGSLIAADVPVAVLAFVEHVHNEVLFIIVGESLNDVAVVLN 226
 QY 194 AIGSFOLDTHNEHAAPFHLGNFLYLFLLSTLLGAATGLISAYVIKKL-YFGRHSTDREVA 252
 Db 227 VFESFVLGDAVTGDCVKGIVSFFVWS-LGCTLVGVIFAPLLSLVTRFTHVRIIEPG 285
 QY 253 LMMLAYSLYMLAELELDLSGILVFFCGIVMHSYTHWNTVETSSRIYTKHTFATLSLAET 312
 Db 286 FVFVLSVSLYLTSEMLSLSLAILAITEFCGICCKQYKVNISEQSAATVRYTMKMLASGAE 345
 QY 313 FIFLYGMDALDDKRSVSDDTPTGTSTAVSSILMGLVMVGVRAAFVPLSFLSLAKKNOS 372
 Db 346 IIFMFLGISAVDPVIV-----TWNTAFVLTLV--FISVTRALGVVLQWILNRYRWQL 398
 QY 373 EKINFNQVVIWMSGLMRGAVSMALAYNKFRAGHTDVRGNAIMITSTITVCLFSTVVF 432
 Db 399 ETID--QVWMSYGG-LRGAVAYALV---VLLDEKKVKEKLFVSTTLIVVFTTIVFQ 450
 QY 433 MLTKPLISYLLPQNATTSMLSDNTPK---STHPLLDQDSFIEPSGNHNVPDPDSIRG 489
 Db 451 LTIKPLVQWLKVKR-----SEQREPKLNEKLHGRAFD-----HILSAIEDISG 493
 QY 490 FLTRPTRTVHY-----WRQFDDSPM 513
 Db 494 -----QIGHNYLRDKWSNEDRKFLSKV 515
 RESULT 5
 NAH5_HUMAN STANDARD; PRT; 896 AA.
 ID NAH5_HUMAN STANDARD; PRT; 896 AA.
 AC Q14940; Q9Y626;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE SODIUM/HYDROGEN EXCHANGER 5 (NA(+)/H(+) EXCHANGER 5) (NHE-5).
 GN SLC9A5 OR NHE5.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;

Query Match 16.8%; Score 462; DB 1; Length 834;
 Best Local Similarity 28.6%; Pred. No. 8.4e-24;
 Matches 146; Conservative 97; Mismatches 203; Indels 64; Gaps 18;

Qy 20 VVALNLFVALLCACIVLGHLEENRWNESITALLIGLGTGVTILLISKG---KSHLLV 76
 Db 57 VIALWILVASLAK---IGFHLSHKVSVPESALLVLG-----LVLGIVWAADHIAS 107

Qy 77 F--SEDFFLYLPPPIFNAGFQVKKOFFNFVTIMLFGAVGTIISCTIISLGTQF 134
 Db 108 FTLPPTVFFVLLPPIVDLAGYFMPNRLFGNLTCTILLVAVGVTVWNAATGSLYGVFL 167

Qy 135 KLDIGTDLG--DYLAIGAIFAATDSVCTLOVLNQ--DETPLYSLVFGGVVNDATSVVV 191
 Db 168 SGLMGDLQIGLDFLFGSLMAAADVPVAVLAVPEEVHVNVEVLIIFVFGESLNDATVVL 227

Qy 192 FNAQSPDLTHLNEAAFHLLGNFLFLSTLLGAATGLISAYVTKKL-YFGRHSDRE 250
 Db 228 YNVESFVALGGDNVTCVDCVKGIIVSEFVVS--LGGVLGVVFAFLLSLVTRTKHVRIE 286

Qy 251 VALMMLMAYLSYMLAEFLDLSGILTFFCGIVMSHYTHNVTTESSRIITKHTFATLSFLA 310
 Db 287 PGFVFIISYLSYLTSEMLSAILATFCGICCKVYKANISEQSATTVRYTMKMLASSA 346

Qy 311 ETFFFLVGMADLIDKWRVSVDTPGTSIAVSSILMGLVMVGEAAFPVPLSLSLNAKN 370
 Db 347 ETIIFMFLGISAVNPFTW-----TWNTAFVLLTLV--FISVYRAIGVLOTWLLNRYRMV 399

Qy 371 QSEKINFNMQVVIWMSGLMRGAVSMALAYNKFTRAGHTDVRGNALIMITITIVVCLFSTVV 430
 Db 400 QLEPID--QVVLSSYGL--RGAAFAFV-----VLLDGDVKVKEKNLRFVSTIIVVFFTVIF 451

Qy 431 PGMLTKPLISYLLPHQNAATSMISDDNTPK---SIHIPLLDQDSFIEPGSNHVPPOSDI 487
 Db 452 QGLTIKPLVOWLVKVR-----SEHREPLNEKLHGRAPD-----HILSAIEDI 494

Qy 488 RGLFTRTRTVHYV---WRQFDDSFMRPV 513
 Db 495 SG-----QIGHNLYRDKWHSFDRKFLSRV 518

RESULT 8
 NAH3_DIDMA STANDARD; PRT; 839 AA.
 AC Q28362;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE SODIUM/HYDROGEN EXCHANGER 3 (NA(+)/H(+) EXCHANGER 3) (NHE-3).
 GN SIC9A3 OR NHE3.
 OS Didelphis marsupialis virginiana (North American opossum).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Metatheria; Didelphimorphia; Didelphidae; Didelphis.
 OX NCBI_taxid=9267;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95358256; PubMed=7631739;
 RA Amemiya M., Yamaji Y., Cano A., Moe O.W., Alpern R.J.;
 RT "Acid incubation increases NHE-3 mRNA abundance in OKP cells.";
 RL Am. J. Physiol. 269:C126-C133(1995).
 CC -!- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED
 CC BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL
 CC CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD
 CC SODIUM ION CHEMICAL GRADIENT. PLAYS AN IMPORTANT ROLE IN SIGNAL
 CC TRANSDUCTION.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -!- PTM: PHOSPHORYLATED (POSSIBLE).
 CC -!- SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.
 CC -!- CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC
 CC DOMAINS IN THE NA(+)/H(+) EXCHANGERS VARY AMONG AUTHORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L42522; AAA98816.1; -;
 CC InterPro; IPR000676; -;
 CC Pfam; PF00999; Na_H_Exchange; 1.
 CC PRINTS; PR01084; NAHEXCHNGR.
 CC PRINTS; PR01087; NAHEXCHNGR3.
 CC Transmembrane; Glycoprotein; Sodium transport; Transport; Symport;
 CC Multigene family; Phosphorylation.
 CC FT DOMAIN 1 13 CYTOPLASMIC (POTENTIAL).
 CC FT TRANSMEM 14 31 M1 (POTENTIAL).
 CC FT DOMAIN 32 65 EXTRACELLULAR (POTENTIAL).
 CC FT TRANSMEM 66 85 M2 (POTENTIAL).
 CC FT DOMAIN 86 87 CYTOPLASMIC (POTENTIAL).
 CC FT TRANSMEM 88 107 M3 (POTENTIAL).
 CC FT DOMAIN 108 113 EXTRACELLULAR (POTENTIAL).
 CC FT TRANSMEM 114 133 M4 (POTENTIAL).
 CC FT DOMAIN 134 150 CYTOPLASMIC (POTENTIAL).
 CC FT TRANSMEM 151 170 M5 (POTENTIAL).
 CC FT DOMAIN 171 186 EXTRACELLULAR (POTENTIAL).
 CC FT TRANSMEM 187 206 M5A (POTENTIAL).
 CC FT DOMAIN 207 215 CYTOPLASMIC (POTENTIAL).
 CC FT TRANSMEM 216 235 M5B (POTENTIAL).
 CC FT DOMAIN 236 255 EXTRACELLULAR (POTENTIAL).
 CC FT TRANSMEM 256 276 M6 (POTENTIAL).
 CC FT DOMAIN 277 299 CYTOPLASMIC (POTENTIAL).
 CC FT TRANSMEM 300 319 M7 (POTENTIAL).
 CC FT DOMAIN 320 347 EXTRACELLULAR (POTENTIAL).
 CC FT TRANSMEM 348 366 M8 (POTENTIAL).
 CC FT DOMAIN 367 370 CYTOPLASMIC (POTENTIAL).
 CC FT TRANSMEM 373 392 M9 (POTENTIAL).
 CC FT DOMAIN 393 442 EXTRACELLULAR (POTENTIAL).
 CC FT TRANSMEM 443 462 M10 (POTENTIAL).
 CC FT DOMAIN 463 839 CYTOPLASMIC (POTENTIAL).
 CC FT CARBOHYD 331 331 N-LINKED (GLCNAC...) (POTENTIAL).
 CC SEQUENCE 839 AA; 94765 MW; F6F9EF778D5DBBB2 CRC64;

Query Match 16.7%; Score 459; DB 1; Length 839;
 Best Local Similarity 30.9%; Pred. No. 1.3e-23;
 Matches 136; Conservative 86; Mismatches 168; Indels 50; Gaps 16;

Qy 20 VVALNLFVALLCACIVLGHLEENRWNESIT-----ALLIGLGTGVTILLISKG---K 70
 Db 62 IIALWILVASLAK--IVFHL-----SHKVSVPESALLVLG-----LILGIVWA 106

Qy 71 SSHLLVF--SEDFFLYLPPPIFNAGFQVKKOFFNFVTIMLFGAVGTIISCTIISLG 128
 Db 107 ADHIASFLLTPTVFFVLLPPIVDLAGYFMPNRLFGNLTCTILLVAVIGTVWNAATFGLS 166

Qy 129 VTQFFKKLDIGTDLG--DYLAIGAIFAATDSVCTLOVLNQ--DETPLYSLVFGGVVND 185
 Db 167 LYGYLVSIGMGDLISGLLDFLFGSLAAADVPVAVLAVPEEVHVNVDLFIIVFGESLND 226

Qy 186 ATSVVFNALQSPDLTHLNEAAFHLLGNFLFLSTLLGAATGLISAYVTKKL-YFGR 244
 Db 227 AVTVVLNVDSFVSLGADKVTGDCVKGIIVSEFVVS--LGGTLIGIIFAFLLSLVTRFK 285

Qy 245 HSTOREVALMMLMAYLSYMLAEFLDLSGILTFFCGIVMSHYTHNVTTESSRIITKHTFA 304
 Db 286 HVRIIEFGFVFIISYLSYLTSEMLSAILATFCGICCKVYKANISEQSATTVRYTMK 345

Qy 305 TLSFLATFFLYVGMADLIDKWRVSVDTPGTSIAVSSILMGLVMVGEAAFPVPLSL 362
 Db 346 MLASGAETIIFMFLGISAVNPFTW-----TWNTAF-----ILLTLVFSIVRAIGVLOTW 396

Qy 363 LSNLAKKNQSEKINFNMQVVIWMSGLMRGAVSMALAYNKFTRAGHTDVRGNALIMITIT 422

```
Db 397 LNKRYMVOLEID---QVMVSGGL-RGAVAYALV-----VLLDEKKVKERNLFWSTII 448
QY 423 VCLFSTVVFVCMGLTKPLISYL 442
Db 449 VVFTVIFQGLTIKPLVQWL 468

RESULT 9
NAH2_RAT
ID NAH2_RAT STANDARD: PRT: 813 AA.
AC P48763; Q16434;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE SODIUM/HYDROGEN EXCHANGER 2 (NA(+)/H(+)) EXCHANGER 2) (NHE-2) (H7).
GN SLC9A2 OR NHE2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A. (LONG ISOFORM).
RC TISSUE=Stomach;
RX MEDLINE=93280160; PubMed=7685026;
RA Wang Z., Orłowski J., Shull G.E.;
RT "Primary structure and functional expression of a novel
RT gastrointestinal isoform of the rat Na/H exchanger."
RL J. Biol. Chem. 268:11925-11928(1993).
RN [2]
RP SEQUENCE FROM N.A. (SHORT ISOFORM).
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Small Intestine;
RX MEDLINE=93248205; PubMed=7683411;
RA Collins J.F., Honda T., Knobel S., Bulus N.M., Conary J.,
RA Dubois R., Ghishan F.K.;
RT "Molecular cloning, sequencing, tissue distribution, and functional
RT expression of a Na+/H+ exchanger (NHE-2).";
RL Proc. Natl. Acad. Sci. U.S.A. 90:3938-3942(1993).
RN [3]
RP SEQUENCE FROM N.A. (SHORT ISOFORM).
RC TISSUE=Liver;
RX MEDLINE=96129297; PubMed=8595899;
RA Ghishan F.K., Knobel S.M., Summar M.;
RT "Molecular cloning, sequencing, chromosomal localization, and tissue
RT distribution of the human Na+/H+ exchanger (SLC9A2).";
RL Genomics 30:25-30(1995).
CC -1- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED
CC BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL
CC CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD
CC SODIUM ION CHEMICAL GRADIENT. SEEMS TO PLAY AN IMPORTANT ROLE IN
CC COLONIC SODIUM ABSORPTION.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: PREDOMINANTLY IN SMALL INTESTINE, COLON, AND
CC STOMACH, WITH MUCH LOWER LEVELS IN SKELETAL MUSCLE, KIDNEY, BRAIN,
CC TESTIS, UTERUS, HEART, AND LUNG.
CC -1- PM: PHOSPHORYLATED (POSSIBLE).
CC -1- SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.
CC -1- CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC
CC DOMAINS IN THE NA(+)/H(+) EXCHANGERS VARY AMONG AUTHORS.
CC -1- CAUTION: REF.3 SEQUENCE WAS ORIGINALLY THOUGHT TO ORIGINATE FROM
CC HUMAN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; L11236; AAA72350.1; -
```

```
DR EMBL; L11004; AAA75406.1; -.
DR EMBL; S81591; AAB36180.1; -.
DR InterPro; IPR000676; -.
DR InterPro; IPR001953; -.
DR Pfam; PF00999; Na_H_Exchange; 1.
DR PRINTS; PRO1084; NAHEXCHNGR.
DR PRINTS; PRO1086; NAHEXCHNGR2.
DR Transmembrane; Glycoprotein; Sodium transport; Transport; Symport;
KW Multigene family; Phosphorylation; ALTERNATIVE SPLICING.
KW Cytoplasmic (POTENTIAL).
FT DOMAIN 1 13
FT A (M1) HYDROPHOBIC.
FT DOMAIN 14 34
FT B (M2) HYDROPHOBIC.
FT DOMAIN 35 80
FT C (M3) (POTENTIAL).
FT DOMAIN 102 107
FT CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 108 128
FT D (M4) (POTENTIAL).
FT DOMAIN 129 139
FT EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 140 160
FT E (M5) (POTENTIAL).
FT DOMAIN 161 169
FT F (M5A) (POTENTIAL).
FT TRANSMEM 170 190
FT G (M5B) (POTENTIAL).
FT DOMAIN 191 209
FT H (M6) (POTENTIAL).
FT TRANSMEM 210 230
FT I (M7) (POTENTIAL).
FT TRANSMEM 231 237
FT J (M8) (POTENTIAL).
FT TRANSMEM 238 258
FT K (M9) (POTENTIAL).
FT DOMAIN 259 278
FT L, HYDROPHOBIC.
FT TRANSMEM 279 299
FT M13 (POTENTIAL).
FT DOMAIN 300 308
FT N-LINKED (GLCNAC...) (POTENTIAL).
FT TRANSMEM 309 329
FT MISSING (IN SHORT ISOFORM).
FT DOMAIN 330 361
FT H -> HW (IN REF. 3).
FT TRANSMEM 362 382
FT LYQIRQ -> SLSNPPA (IN REF. 3).
FT TRANSMEM 383 392
FT A -> P (IN REF. 3).
FT TRANSMEM 393 413
FT V -> G (IN REF. 2).
FT DOMAIN 414 430
FT CONFLICT 504 504
FT TRANSMEM 431 451
FT CONFLICT 610 616
FT DOMAIN 452 459
FT CONFLICT 742 742
FT TRANSMEM 480 480
FT CONFLICT 786 786
FT TRANSMEM 481 813
FT SEQUENCE 813 AA; 29727267D7085845 CRC64;
SQ
Query Match 16.5%; Score 455.5; DB 1; Length 813;
Best Local Similarity 28.1%; Pred. No. 2.2e-23;
Matches 149; Conservative 94; Mismatches 206; Indels 81; Gaps 22;
QY 7 SKLPSTSDHASV---VALNLFVALLCACIVLGHLLLE-NRWMNESITALLIGLGTGT 62
Db 67 SRLP-VFTLDYPHVQIPFEITLWILLASLAKIGFHLHYHKLPTIVPESCLLIMVGLLGGI 125
QY 63 ILLISKGKSHLLVSEDLFFIYLLPPIIFNAGFOVKKQFFNFVITMIFGAVGTIISC 122
Db 126 IFGVDE-KSPAM--KTDVFFLLPPIVLDAGYPMPTPFENIGTIFWVAVVGLWN- 181
QY 123 TTIISLV-TQFFKKLDIGTDLGDY-----LAIGAIFAATDSVCTQLVL-NODETPLYIS 175
Db 182 ---SIGIGLSLFGICQIEAFGLSDITLLQNLIFGLSLISAVDPVAVLAVFENIHVNEQYI 238
QY 176 LVFGEVNDATSVVFNATQSF-DLTHLNHEAAFHLLGNFLYLLSTLLGAATGLISA 234
Db 239 LVFGEGLNDATVVVLYNLFNFSFQCKTIQVDVDFAGTANFVVGIGVIGLIGLGTAA 298
QY 235 YVKKLYLFGHRHSTDEVALMMLMAYLSYMLAEFLDLSGILTFFCGIVMNSHYTHNVTES 294
Db 299 FTR--FTHNIRVIEPLFVFLYSVLYTAEFMHLSGIMATTACAMTNKYVEENVSQK 355
QY 295 SRITTKHIFATLSFAEFIFLYVGMALDID--KWSVSDTTPGTSIAVSLMLGLVMV 351
Db 356 SYTTIKYFMKMLSSVSETLIFIFMGVSTVGRKNHWNFAVCFPT-----LAFCLI 404
```

QY 352 GRAAFVPLSLAKKQSEKINFMQVVIWWSGLMRGAVSMALAY-----NKETRAGH 407
Db 405 WRALGVFLTVQVINWFR---TIPLTFKQDFIITAYGGL-RGAICFALVFLPATVFPFR--- 457
QY 408 TDVRGNAMITSTITVCLFSTVVFGLMTPLISYL-LPHQNAITTSMLSDNDTPKSHIPL 466
Db 458 -----KKLFIITAIVVIFFTFGLTIRPLVEFLDVKRSNKQQAQVSEE-----IHCRRF 507
QY 467 LDQ-DSFIEPSGNHNVPRDPSIRGFLTRPTVHYWYR-----QFDDSEMR 511
Db 508 FDHVKTGIEDVCGH-----WGHNFWRDKFKFDKDKYL 540
RESULT 10
NAH2_RABIT STANDARD; PRT; 809 AA.
AC P50482;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE SODIUM/HYDROGEN EXCHANGER 2 (NA(+)/H(+)) EXCHANGER 2) (NHE-2).
GN SLC9A2 OR NHE2.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEW ZEALAND WHITE;
RX MEDLINE=93280159; PubMed=7685025;
RA Tse C.-M., Levine S.A., Yun C.H., Montrose M.H., Little P.J.,
RA Pouyssegur J., Donowitz M.;
RT "Cloning and expression of a rabbit cDNA encoding a serum-activated
RT ethylisopropylamide-resistant epithelial Na+/H+ exchanger isoform
RT (NHE-2).";
RL J. Biol. Chem. 268:11917-11924(1993).
CC -!- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED
CC BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL
CC CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD
CC SODIUM ION CHEMICAL GRADIENT. SEEMS TO PLAY AN IMPORTANT ROLE IN
CC COLONIC SODIUM ABSORPTION
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: HIGH LEVELS IN INTESTINE AND KIDNEY.
CC -!- PTM: PHOSPHORYLATED (POSSIBLE).
CC -!- SIMILARITY: BELONGS TO THE NA(+)/H(+)) EXCHANGER FAMILY.
CC -!- CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC
CC DOMAINS IN THE NA(+)/H(+)) EXCHANGERS VARY AMONG AUTHORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L13733; ; NOT_ANNOTATED_CDS.
DR InterPro; IPR000676; .
DR InterPro; IPR001953; .
DR Pfam; PF00999; Na.H.Exchanger; 1.
DR PRINTS; PR01084; NAHEXCHNGR.
DR PRINTS; PR01086; NAHEXCHNGR2.
KW Transmembrane; Glycoprotein; Sodium transport; Transport; Symport;
KW Multigene family; Phosphorylation.
FT DOMAIN 1 16 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 17 37 A (M1) (POTENTIAL).
FT DOMAIN 38 79 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 80 100 B (M2) (POTENTIAL).
FT DOMAIN 101 106 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 107 127 C (M3) (POTENTIAL).
FT DOMAIN 128 138 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 139 159 D (M4) (POTENTIAL).
FT DOMAIN 160 168 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 169 189 E (M5) (POTENTIAL).
FT DOMAIN 190 208 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 209 229 F (M5A) (POTENTIAL).
FT DOMAIN 230 236 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 237 257 G (M5B) (POTENTIAL).
FT DOMAIN 258 277 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 278 298 H (M6) (POTENTIAL).
FT DOMAIN 299 307 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 308 328 I (M7) (POTENTIAL).
FT DOMAIN 329 360 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 361 381 J (M8) (POTENTIAL).
FT DOMAIN 382 391 K (M9) (POTENTIAL).
FT TRANSMEM 392 412 L (M10) (POTENTIAL).
FT DOMAIN 413 429 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 430 450 M (M11) (POTENTIAL).
FT DOMAIN 451 458 N-LINKED (GLCNAC... (POTENTIAL).
FT TRANSMEM 459 479
FT DOMAIN 480 809
FT CARBOHYD 350 350
SQ SEQUENCE 809 AA; 90744 MW; DBD00B45443D87A6 CRC64;

Query Match 16.5%; Score 454.5; DB 1; Length 809;
Best Local Similarity 28.4%; Pred. No. 2.6e-23;
Matches 150; Conservative 94; Mismatches 206; Indels 79; Gaps 22;

QY 7 SKLPSLSTSDHASV---VALNLFVALLCACIVLGHLEE-NRWNNESITALLIGLGTGVT 62
Db 66 SRLP-VFTLDYPHQVQIPFEITLWILLASLAKIGFLYHKLPTIVPESCLLMVGLLGGI 124
QY 63 ILLSKGSSHLVFSDELFFIYLLPPIFNAGFOVKKKOFFRNFVTIMLFGAVGTIISC 122
Db 125 IFGVDE-KSPAM--KTVDFVLYLLPVIDDAGFMFTRFFEMLGTIFWVAVGTUWS 181
QY 123 TIISLGVTQFKLDIGTDFLDGY-----LAIGAIFAATDSVCTQLVL-NODETPLYSL 176
Db 182 --IGIGVS-LFGICQIEAFGLSDITLLQNLFLGSLISAVDPVAVLAVFENHVNQYL 238
QY 177 VFGGVNDATSVVFNAIQSF-DLTHLNEAFAHLGNFLYLLFLLSTLLGAATGLISAY 235
Db 239 VFGESLNDATVVVLYNLFSQCMKTETIDVFAGIANFVVGIGVGLIGLGFIAAF 298
QY 236 VIKKLYGRHSTDREVALMMLMAYLSYMLAEFLDLSLAVFCGIVMHSYTHVHNTESS 295
Db 299 TTR---ETHNIRVIEPLFVLYSLVSYTAEMPHLSGIMAITACAMTNKYVENVSQKS 355
QY 296 RITTKHFTATLSLAETFIIFYGMDALDID---KWRVSVDTPGTSTAVSSILMGLVMVG 352
Db 356 YTTIKYFMKMLSSVSETLIETFMGVSTVGKNHWNNAFVCF-----LAFCLIW 404
QY 353 RAAVFPPLSLSLAKKQSEKINFMQVVIWWSGLMRGAVSMALAY-----NKETRAGHT 408
Db 405 RALGVFLTVQVINWFR---TIPLTFKQDFIITAYGGL-RGAICFALVFLPATVFPFR--- 456
QY 409 DVRGNAIMTSTITVCLFSTVVFGLMTPLISYL-LPHQNAITTSMLSDNDTPKSHIPL 467
Db 457 ----KKLFIITAIVVIFFTFGLTIRPLVEFLDVKRSNKQQAQVSEE-----IHCRRF 507
QY 468 DQ-DSFIEPSGNHNVPRDPSIRGFLTRPTVHYWYR-----QFDDSEMR 511
Db 508 DHVKTGIEDVCGH-----WGHNFWRDKFKFDKDKYL 539

RESULT 11
NAH2_HUMAN STANDARD; PRT; 812 AA.
ID NAH2_HUMAN AC Q9UBX0;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE SODIUM/HYDROGEN EXCHANGER 2 (NA(+)/H(+)) EXCHANGER 2) (NHE-2).
GN SLC9A2 OR NHE2.
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RX MEDLINE=99375108; PubMed=10444453;
RA Malakooti J., Dahdal R.Y., Schmidt L., Layden T.J., Dudeja P.K.,
RA Ramaswamy K.;
RT "Molecular cloning, tissue distribution, and functional expression of
RT the human Na(+)/H(+) exchanger NHE2.";
RL Am. J. Physiol. 277:G383-G390(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Hou S., Wohldmann P.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED
CC BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL
CC CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD
CC SODIUM ION CHEMICAL GRADIENT. SEEMS TO PLAY AN IMPORTANT ROLE IN
CC COLONIC SODIUM ABSORPTION.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN SKELETAL MUSCLE, COLON AND
CC KIDNEY. LOWER LEVELS IN THE TESTIS, PROSTATE, OVARY, AND SMALL
CC INTESTINE.
CC -!- PTM: PHOSPHORYLATED (POSSIBLE).
CC -!- SIMILARITY: BELONGS TO THE Na(+)/H(+) EXCHANGER FAMILY.
CC -!- CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC
CC DOMAINS IN THE Na(+)/H(+) EXCHANGERS VARY AMONG AUTHORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
ER EMBL; AF073299; AAD41635.1; -;
DR EMBL; AC007239; AAF19248.1; -;
DR MIM; 600530; -;
DR InterPro; IPR000676; -;
DR InterPro; IPR001953; -;
DR Pfam; PF00999; Na_H_Exchange; 1.
DR PRINTS; PR01084; NAHEXCHNGR.
DR PRINTS; PR01086; NAHEXCHNGR2.
KW Transmembrane; Glycoprotein; Sodium transport; Transport; Symport;
KW Multigene family; Phosphorylation.
FT DOMAIN 1 13
FT DOMAIN 14 33 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 34 79 A (M1) HYDROPHOBIC.
FT DOMAIN 80 100 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 101 106 B (M2) HYDROPHOBIC.
FT DOMAIN 107 127 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 128 138 C (M3) (POTENTIAL).
FT DOMAIN 139 159 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 160 168 D (M4) (POTENTIAL).
FT TRANSMEM 169 189 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 190 208 E (M5) (POTENTIAL).
FT DOMAIN 209 229 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 230 236 F (M5A) (POTENTIAL).
FT TRANSMEM 237 257 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 258 277 G (M5B) (POTENTIAL).
FT TRANSMEM 278 298 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 299 307 H (M6) (POTENTIAL).
FT TRANSMEM 308 328 I (M7) (POTENTIAL).
FT DOMAIN 329 360 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 361 381 J (M8) (POTENTIAL).
FT DOMAIN 382 391 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 392 412 K (M9) (POTENTIAL).
FT DOMAIN 413 429 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 430 450 L, HYDROPHOBIC.
FT DOMAIN 451 458 EXTRACELLULAR (POTENTIAL).
FT DOMAIN

FT TRANSMEM 459 479 M13 (POTENTIAL).
FT DOMAIN 480 812 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 350 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 812 AA; 91519 MW; 17BE177DC3830D0A CRC64;

Query Match 16.4%; Score 452.5; DB 1; Length 812;
Best Local Similarity 28.4%; Pred. No. 3.5e-23;
Matches 150; Conservative 96; Mismatches 204; Indels 79; Gaps 22;

QY 7 SKLPSTSDSHASV---VALNLFVALLCACIVLGHLLER-NRWMNESITALLIGLGTGT 62
DB 66 SRLP-VFTLDYPHVQPPFEITLWILLASLAKGFHLYHLPTIVPESCLLIWGLLGI 124

QY 63 ILLISKSKSHLLVSEDLFFIYLLPPIIFNAGQVKKQKQFFNFVTIMLFGAVGTIISC 122
DB 125 IFGVDE-KSPAPM-KTDVEFLYLLPPIVDAGYFMPTRPFENIGTIPWYAVVCTLWS 181

QY 123 TITSLGVTOFFKKLDIGTFDLGDY-----LATGAFAAATDSVCTLOVL-NQDETPLLISL 176
DB 182 --IGIGVS-LFGICQIEAFGLSDITLLQNLFGSLISAVDPVAVLAVFENIHVNEQLYL 238

QY 177 VEGGVNDATSVVFNATQSF-DLTHLNHEAAFHLLGNFLYFLFLLSTLLGAATGLISAY 235
DB 239 VFGESLNDATVVLNLFKSCQMTIETIDVFAGIANFVVGIGVGLIGLIFLGFIAAF 298

QY 236 VIRKLYFGRHSTDREVALMMLMAYLSMLAEFLDLSGILTFFCGIVMGSHTYHNNVTSS 295
DB 299 TTR--FTHNIRVIEPLFVFLYSYLSYITAEMHLSNAITACAMTMNKYVEENVSKS 355

QY 296 RITTKITFATLSPLAETFTFLYVGMALDID---KWRVSVDTPGTSIAVSSILMLGVWVG 352
DB 356 YTIKIKYFMKMLSSVSETLIFIMGVSTGVKNHEWNMAFVCF-----LAFCLMW 404

QY 353 RAAFPVPLSLNLAKKQSEKINENMVIMWVSGILMRGAVSMALAY----NKETRAGHT 408
DB 405 RALGVFLVQVIN--RFRITPTFKDQIIAYGGL-RGAICFALVFLPAAVFP----- 456

QY 409 DVGRNATITSTITVCLFSTVFEGLTKPLISYL-LPHQNATTSMLSDNTPKSIHPLL 467
DB 457 ---KKLFTAAIWIFFTVILGITIRPLVEFLDKRNSKKQAVSEE-----IYCRLF 507

QY 468 DQ-DSIEPSGNVPRPDSIRGFLRPTVHYWYR---QFDDSFMR 511
DB 508 DHVKTGIEDVCGH-----WGHNFWDRKFKFPDDKYL 539

RESULT 12
NAHL_HUMAN STANDARD; PRT; 815 AA.
AC P19634;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE SODIUM/HYDROGEN EXCHANGER 1 (NA(+)/H(+)) EXCHANGER 1) (NHE-1) (NA+/H+
DE ANTIporter, AMILORIDE-SENSITIVE) (APNH).
GN SLC9A1 OR NHE1 OR APNH1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=89106219; PubMed=2536298;
RA Sargent C., Franchi A., Pouyssegur J.;
RT "Molecular cloning, primary structure, and expression of the human
RL growth factor-activatable Na+/H+ antiporter.";
RN Cell 56:271-280(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=90140739; PubMed=2154036;


```
Db 437 N---KFRIVLTKPQDFIIAYGGL- RGAIAFSLGYLLDKKHPMCD-----LFLTAI 484
Qy 421 ITVCLFSTVFGMLTPLSYLLPHONATTSMLSDNTPKSHIPLDDQDSFIEPSGNHN 480
Db 485 ITVFFTVFVQGMTRPLVDLLAVKKQKQTKR-----SINEEHTQFLD-----HL 530
Qy 481 VPRPDSIRGFLTRPTRTVHYWYR-----QFDDSFMRPVF----- 514
Db 531 LTGIEDICGHYG-----HHWKDKLNRENKKYVKCKLIAGERSKEPQLIAFYHKMEMQ 584
Qy 515 -----GGRGFVPFVPGS-----PTERNPPDLK 537
Db 585 AIELVESGGMGKIPSAVSTVSMQNIHPKALPAERILPALSK 625

RESULT 14
NAHL_PIG
ID NAHL_PIG STANDARD; PRT; 818 AA.
AC P48762;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE SODIUM/HYDROGEN EXCHANGER 1 (NA(+)/H(+)) EXCHANGER 1) (NHE-1).
GN SLC9A1 OR NHE1.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=92087905; PubMed=1661081;
RX Reilly R.F., Hilgendorf F., Blemesderfer D., Sargent C.,
RA Pouyssegur J., Aronson P.S., Slayman C.W., Igarashi P.;
RT "cDNA cloning and immunolocalization of a Na(+)-H+ exchanger in
RT LLC-PK1 renal epithelial cells.";
RL Am. J. Physiol. 261:F1088-F1094(1991).
CC -!- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED
CC BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL
CC CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD
CC SODIUM ION CHEMICAL GRADIENT. PLAYS AN IMPORTANT ROLE IN SIGNAL
CC TRANSDUCTION.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- PTM: PHOSPHORYLATED (POSSIBLE).
CC -!- SIMILARITY: BELONGS TO THE NA(+)/H(+)) EXCHANGER FAMILY.
CC -!- CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC
CC DOMAINS IN THE NA(+)/H(+)) EXCHANGERS VARY AMONG AUTHORS.
CC -!- CAUTION: HYDROPHOBIC DOMAINS A, B AND L ARE NOT BELIEVED TO BE
CC TRANSMEMBRANAL, BUT ONLY MEMBRANE-ASSOCIATED.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M89631; AAA31092.1; -
DR EMBL; S71135; AAB20633.1; -
DR InterPro; IPR000576; -
DR InterPro; IPR001970; -
DR Pfam; PF00999; Na_H-Exchanger; 1.
DR PRINTS; PR01084; NAHEXCHNGR.
DR PRINTS; PR01085; NAHEXCHNGR.
KW Transmembrane; Glycoprotein; Sodium transport; Transport; Symport;
KW Multigene family; Phosphorylation.
FT DOMAIN 1 11 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 12 31 A (M1) HYDROPHOBIC.
FT DOMAIN 32 101 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 102 123 B (M2) HYDROPHOBIC.
FT DOMAIN 124 126 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 127 146 C (M3) (POTENTIAL).
```

```
FT DOMAIN 147 158 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 159 179 D (M4) (POTENTIAL).
FT DOMAIN 180 184 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 185 206 E (M5) (POTENTIAL).
FT DOMAIN 207 226 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 227 247 F (M5A) (POTENTIAL).
FT DOMAIN 248 256 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 257 278 G (M5B) (POTENTIAL).
FT DOMAIN 279 297 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 298 318 H (M6) (POTENTIAL).
FT DOMAIN 319 332 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 333 353 I (M7) (POTENTIAL).
FT DOMAIN 354 384 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 385 406 J (M8) (POTENTIAL).
FT DOMAIN 407 412 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 413 434 K (M9) (POTENTIAL).
FT DOMAIN 435 448 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 449 469 L, HYDROPHOBIC.
FT DOMAIN 470 478 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 479 499 M (M10) (POTENTIAL).
FT DOMAIN 500 818 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 370 370 N-LINKED (GLCNAC...) (POTENTIAL).
FT CONFLICT 683 683 H -> Y (IN AAB20633).
SQ SEQUENCE 818 AA; 90987 MW; 9329F7D9A51D3DC9 CRC64;

Query Match 15.8%; Score 436; DB 1; Length 818;
Best Local Similarity 25.5%; Pred. No. 4.5e-22;
Matches 148; Conservative 105; Mismatches 202; Indels 126; Gaps 22;

Qy 23 LNLFWALLCACIVLGHLLLE-NRWNNESITALLGLGTGVTILLISKSKSHLLVFSDEL 81
Db 105 ISLWILLACLKMGIFHVPTISSIPESCLLVGLVGLIKAV--GETPPFL--OSEV 160
Qy 82 FFYLLPPIIPNAGFOVKKQFFNFVTIMLFAGVGTI-----ISCTIISLGVTVQFF 133
Db 161 FFELLPPIILDAGYPLPQFTENLGTILFAVVGTLWNAFFLGLMYAVCLVGGQ-- 218
Qy 134 KKLDTGFDLGDYLAIGAFAAFTASVCTQLVNODE--TPLLYSLVFGEGVVDNATSVVF 192
Db 219 ----INNIGLDNLFGSIISAVDPVAVLFAVEIHNELLHLVFGESLINDAVTVLY 274
Qy 193 NAIQSFDLTHLNHEAFAFLHGLNFLYFLLS---TLIGAATGLISAYVIKKLYFGRHSTD 249
Db 275 HLFEF--ANYDRVGVIVDLVGLFSLFVSVGLGVGVGVIAAFTSR---FTSHIRVI 329
Qy 250 EVALMLMAYLSYMLAEFLDLSGILTVFVFCGIVMSHYTHWNTESSRITTKHTFATLSFL 309
Db 330 EPLEFVLYSYMAYLSAEFLHLSGIMALTASGVVMRPYVEANISHKSHTTIKYFLKMSSV 389
Qy 310 AETFIELVGMADALDIDKRSVSDTPCT----SIAYSSILMGLVMVYGRAAFVPLSFLS 364
Db 390 SETLIFILFG-----VSTVAGSHHWNWTFVISTILL--FCLIARVLGLVGLTWFI 436
Qy 365 NLAKKNQSEKINENMQVIVMWSGLMRGAVSMALAY----NKFTRAGHTDVRGNAIMTST 420
Db 437 N---KERIVLTKPQDFIIAYGGL- RGAIAFSLGHLDDKNHPPMCD-----LFLTAI 484
Qy 421 ITVCLFSTVFGMLTPLSYLLPHONATTSMLSDNTPKSHIPLDDQDSFIEPSGNHN 480
Db 485 ITVFFTVFVQGMTRPLVDLLAVKKQKQTKR-----SINEEHTQFLD-----HL 530
Qy 481 VPRPDSIRGFLTRPTRTVHYWYR-----QFDDSFMRPVF----- 514
Db 531 LTGIEDICGHYG-----HHWKDKLNRENKKYVKCKLIAGERSKEPQLIAFYHKMEMQ 584
Qy 515 -----GGRGFVPFVPGS-----PTERNPPDLK 537
Db 585 AIELVESGGMGKIPSAVSTVSMQNIHPKALPAERILPALSK 625

RESULT 15
NAHL_ONCMY
```

ID	NAHB_ONCMY	STANDARD;	PRT;	759 AA.
DT	Q01345;			
DT	01-JUN-1994 (Rel. 29, Created)			
DT	01-JUN-1994 (Rel. 29, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	NA(+) /H(+) EXCHANGER BETA (NA(+)/H(+)) ANTIPOTER (BETA-NHE).			
OS	Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei;			
OC	Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.			
NCBI_TaxID=8022;				
ON	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Nucleated erythrocyte;			
RX	MEDLINE=92357712; PubMed=1379718;			
RT	Borgese F., Sardet C., Cappadoro M., Pouyssegur J., Motaïs R.;			
RT	"Cloning and expression of a CAMP-activated Na+/H+ exchanger;			
RT	evidence that the cytoplasmic domain mediates hormonal regulation.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 89:6765-6769(1992).			
CC	-1- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD SODIUM ION CHEMICAL GRADIENT.			
CC	-1- SUBCELLULAR LOCATION: BASOLATERAL MEMBRANE.			
CC	-1- PTM: ACTIVATED BY CAMP, PROTEIN KINASE A AND PROTEIN KINASE C.			
CC	-1- MISCELLANEOUS: INHIBITED BY AMILORIDE AND 5-AMINO-SUBSTITUTED DERIVATIVES AND ACTIVATED IN A COOPERATIVE FASHION BY INTRACELLULAR H+.			
CC	-1- SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.			
CC	-----			
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch).			
CC	-----			
EMBL; M94581; AAA49549.1; --				
DR	InterPro: IPR000676; --			
DR	Pfam: PF00999; Na_H_Exchange; 1.			
DR	PRINTS; PR01084; NAHXCXNHR.			
KW	Transmembrane; Glycoprotein; Sodium transport; Transport; Symport;			
KW	Mitogene family; Phosphorylation.			
FT	DOMAIN 1 14			
FT	TRANSMEM 15 34			
FT	DOMAIN 35 75			
FT	TRANSMEM 76 95			
FT	DOMAIN 96 97			
FT	TRANSMEM 98 117			
FT	DOMAIN 118 122			
FT	TRANSMEM 123 142			
FT	DOMAIN 143 149			
FT	TRANSMEM 150 169			
FT	DOMAIN 170 195			
FT	TRANSMEM 196 214			
FT	DOMAIN 215 225			
FT	TRANSMEM 226 244			
FT	DOMAIN 245 261			
FT	TRANSMEM 262 282			
FT	DOMAIN 283 311			
FT	TRANSMEM 312 330			
FT	DOMAIN 331 352			
FT	TRANSMEM 353 372			
FT	DOMAIN 373 376			
FT	TRANSMEM 377 398			
FT	DOMAIN 399 446			
FT	TRANSMEM 447 467			
FT	DOMAIN 468 759			
FT	MOD_RES 641 641			
FT	MOD_RES 648 648			
FT	CARBOHYD 49 49			
FT	CARBOHYD 338 338			

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run On: November 18, 2001, 08:47:25 ; Search time 2907.92 Seconds
(without alignments)
11585.173 Million cell updates/sec

Title: US-09-271-584A-1
Perfect score: 2178
Sequence: 1 cctctctgttcttctctg.....aaaaaaaaaaaaaaaaaaaaa 2178

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:
1: gb_ba1.*
2: gb_ba2.*
3: gb_ba3.*
4: gb_in1.*
5: gb_in2.*
6: gb_in3.*
7: gb_om.*
8: gb_ov.*
9: gb_pat1.*
10: gb_pat2.*
11: gb_ph.*
12: gb_p11.*
13: gb_p12.*
14: gb_p13.*
15: gb_p14.*
16: em_ba1.*
17: em_ba2.*
18: em_fun.*
19: em_htgo_hum.*
20: em_htgo_inv.*
21: em_htgo_rod.*
22: em_htg_hum1.*
23: em_htg_hum2.*
24: em_htg_hum3.*
25: em_htg_hum4.*
26: em_htg_hum5.*
27: em_htg_hum6.*
28: em_htg_hum7.*
29: em_htg_hum8.*
30: em_htg_inv1.*
31: em_htg_inv2.*
32: em_htg_other.*
33: em_htg_rod.*
34: em_hum1.*
35: em_hum2.*
36: em_hum3.*
37: em_hum4.*
38: em_hum5.*
39: em_hum6.*
40: em_hum7.*
41: em_in.*
42: em_om.*
43: em_or.*

44: em_ov.*
45: em_pat.*
46: em_ph.*
47: em_pi.*
48: em_ro.*
49: em_sts.*
50: em_sy.*
51: em_un.*
52: em_vi.*
53: gb_sts1.*
54: gb_sts2.*
55: gb_sts3.*
56: gb_sy.*
57: gb_un.*
58: gb_vil.*
59: gb_v12.*
60: gb_htg1.*
61: gb_htg2.*
62: gb_htg3.*
63: gb_htg4.*
64: gb_htg5.*
65: gb_htg6.*
66: gb_htg7.*
67: gb_htg8.*
68: gb_htg9.*
69: gb_htg10.*
70: gb_htg11.*
71: gb_htg12.*
72: gb_htg13.*
73: gb_htg14.*
74: gb_htg15.*
75: gb_htg16.*
76: gb_htg17.*
77: gb_htg18.*
78: gb_htg19.*
79: gb_htg20.*
80: gb_htg21.*
81: gb_htg22.*
82: gb_htg23.*
83: gb_htg24.*
84: gb_htg25.*
85: gb_pr1.*
86: gb_pr2.*
87: gb_pr3.*
88: gb_pr4.*
89: gb_pr5.*
90: gb_pr6.*
91: gb_pr7.*
92: gb_pr8.*
93: gb_pr9.*
94: gb_ro1.*
95: gb_ro2.*
96: gb_in4.*
97: gb_pr10.*
98: em_ba3.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1615.8	74.2	1619	12	AF056190	Arabidops
2	1614	74.1	1614	12	AF106324	Arabidops
3	832.8	38.2	2637	12	AB038492	Atriplex
4	774.8	35.6	2218	12	AB033989	Ipomoea n
5	768.8	35.3	2313	12	AB021878	Oryza sat
6	693	31.8	90019	15	TM021B04	Arabidops
7	266.6	12.2	600	14	AY028416	Citrus x
8	229.6	10.5	421	13	AF279671	Mesembrya


```
Db 721 TATACTTTGGAAGCACTCAACTGACGAGAGGTTGCCCTTATGATGCTTATGGCGTATC 780
Qy 1064 ttcttatctgtcgtgagccttttcgaacttgagcgtatccctcaactgtgtttttctgtg 1123
Db 781 TTTCTTATATGCTTCTGAGCTTTTCACATGAGCGGTATCCACATGCTTTTCTGTG 840
Qy 1124 gtattgtatgtccattacacatggtgcacaaatgtaaacgagagctcaagaataacaacaa 1183
Db 841 GTATTGTGATGCCAATTACACATGGCACAAATGTAACGGAGAGCTCAAGAATAACACAA 900
Qy 1184 agcatacctttgcaactttgttcattttcttcgagagacatttatttctgtatgtggaa 1243
Db 901 AGCATACCTTTGCAACTTTTGCACTTTCTCGGAGACATTTATTTCTGTATGTGGAA 960
Qy 1244 tggatgcttggacattgacaagtggagatccgtgagtgacacaccggaacatcatcg 1303
Db 961 TGGATGCTTTGGACATTTGACAAGTGGAGATCCGTGAGTGACACACCGGGAACATCGATCG 1020
Qy 1304 cagtgaagctcaactcaatgggttggtgctatgagagtgctgtatctatgctgtcgttccgt 1363
Db 1021 CAGTGAAGCTCAATCCTAATGGGTGCTGTCATGCTTGAAGAGCAGCGTTCGTCTTTCCGT 1080
Qy 1364 tategtttctataacttagccaagaagaatacaaacgagagaaatacaactttaacatgc 1423
Db 1081 TATCGTTTCTATCTAACTTAGCCAAGAAGAAATCAAGCCGAGAAATCAACTTTAATCATGC 1140
Qy 1424 aggttgattgttggtgctgctcatgagagtgctgtatctatgctgtcgttcttgataca 1483
Db 1141 AGGTTGFGATTTGGTGTGCTGGTCTCATGAGAGTGCTGTATCTATGGCTCTTGCATACA 1200
Qy 1484 acaagtttaacaggccgggacacacagatgtagcggggaatgcaatcatgatacagagta 1543
Db 1201 ACAAGTTTACAAGGGCCGGSCACACAGATGTACCGGGAATGCAATCATGATCACAGTA 1260
Qy 1544 cgataactgtctgtcttttagcacagtgtgttgtagtgatgctgacaaacacactataa 1603
Db 1261 CGATACTGTCTCTTTTGTAGCACAGTGGTGTGGTATGCTGACCAACCACTCATATA 1320
Qy 1604 gctactattacgcaccagaaacgcccacacagcagcagtgatctatgtagcaaacaccccaa 1663
Db 1321 GCTACCTATTACCGCACCAAGAGCCACCAAGCAGCATGTTATCTGATGACACACCCCAA 1380
Qy 1664 aatccatacatcctttgttgaccacaagactcgttcaattgagccttcagggaaccaca 1723
Db 1381 AATCCATACATATCCCTTTTGTGGACCAAGACTCGCTCATTTAGCGCTTCAGGGAACCA 1440
Qy 1724 atgtgctcggcgtgacagtatacgtggtctcttgacacgcccactcgaacccgtgcatt 1783
Db 1441 ATGTGCTCGGCTGACATGATACGTGGCTTCTTGACACGGCCCACTCGAACCGTGCAAT 1500
Qy 1784 actactggagacaatttgatgactccttcacgaccccgcttcttgagggtcgtggtttg 1843
Db 1501 ACTACTGGAGACAATTTGATGACTCTTTTCATGCGACCCGCTCTTGGAGGTCTGGCTTG 1560
Qy 1844 taccctttgtcaggttctccaactgagagaaacccctcctgatatcttagtaaggttga 1902
Db 1561 TACCCTTTGTCCAGGTTCTCCAAGTGTAGAGAAACCCCTCTGATCTTAGTAAGGCTTGA 1619

RESULT 2
AF106324 1614 bp mRNA PLN 03-MAR-1999
LOCUS Arabidopsis thaliana sodium proton exchanger Nhxl mRNA, partial
DEFINITION cds.
ACCESSION AF106324
VERSION AF106324.1 GI:4324596
KEYWORDS thale cress.
SOURCE Arabidopsis.thaliana
ORGANISM Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 1614)
```

```
AUTHORS Gaxiola,R.A., Rao,R., Sherman,A., Grisafi,P., Alper,S.L. and
Fink,G.R.
TITLE The Arabidopsis thaliana proton transporters, AtNhxl and Avp1, can
function in cation detoxification in yeast
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (4), 1480-1485 (1999)
MEDLINE 99145575
REFERENCE 2 (bases 1 to 1614)
AUTHORS Gaxiola,R.A., Rao,R., Sherman,A., Grisafi,P., Alper,S.L. and
Fink,G.R.
TITLE Direct Submission
JOURNAL Submitted (12-NOV-1998) Whitehead, Nine Cambridge Center,
Cambridge, MA 02142, USA
FEATURES
    source
        1..1614
            /organism="Arabidopsis thaliana"
            /cultivar="Columbia"
            /db_xref="taxon:3702"
            1..>1614
            /codon_start=1
            /product="sodium proton exchanger Nhxl"
            /protein_id="AAD16946.1"
            /translation="MLDSLVSKLPSLSTSDHASVVALNLFVALICACIVLHLEENR
WMNESITALLIGLGTVTIILLISKGKSHLLVFSEDLFFIYLLPPIIFNAGRFVKKQ
FFRNFVTIMLFGAVGTIISCTIISLGVTFQFKLDIGTDFLDYLAIGAFATDSVC
TLQVLNODETPLLISLVFGGVYNDATSVVFNAIQSFDLTHLNHEAFAHLLGNFLYL
FLSLTLLAGATGLISAVIKKLYFGRHSDREVALMMLMAYLSYMAELDFLGSILTV
FFCGIVNSHYTHNVTESRITTKHTFATISFLAETIFLYVGMDALDIDKRSVSDT
PGTSIAVSSILMLGVMGVAAFPVPLSFLSLAKNQSKEINFNMQVVIWWSGLMRGA
VSMALAYNKFTFRAGTIDVRGNAIMTSTITVCLFSTVVGMLTKPLISLYLLPHONATT
SMLSDNTPKSIHIPLLDQSFTEPPSNHNVPDPDSIRGLFTRTRIVHYWRFDDSD
EMRPVFGRGFVFPVGPSPTERNPPDLKA"
BASE COUNT 388 a 346 c 358 g 522 t
ORIGIN
Query Match 74.1%; Score 1614; DB 12; Length 1614;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1614; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 286 atgttggaattctctagtgatgaaactgccttcgttatcgacatctgtacacgcttctgtg 345
Db 1 ATGTGGATTTCTAGTGTGAAACTGCCTTCGTTATCGACATCTGATCAGCTTCTGTG 60
Qy 346 gtccgttgaaactcctttgtgcaactcttctgtcgtgtattgttctgttcattctttg 405
Db 61 GTTGGTTGAATCTCTTTGTGCACCTCTTTGTGCTGTATTGTCTTGTGTCATCTTTTG 120
Qy 406 gaagagaatagatggatgaacgaatcccatcacccctctgttgattgggctaggcactggt 465
Db 121 GAAGAGAATAGATGATGATGAACGAATCCATCACCGCTTGTGATTTGGCTAGGCACATG 180
Qy 466 gttaccattttgttgattagtaagggaaaaagctcgcattctctcgttcttagtgaagat 525
Db 181 GTTACCATTTTGTGTGATTAGTAAAGGAAAAAGCTCGCATCTCTCGTCTTTAGTGAAGAT 240
Qy 526 cttctctcatatatcttttgccaccattatattcaatcagaggtttcaagtaaaaaag 585
Db 241 CTTTCTTCATATATCTTTTGGCCACCCCATATATTCATTCAGTCAGGGTTTCAAGTAAAAAG 300
Qy 586 aagcagtttttcccgaatttctgtaactattatgctttttgtgctgttgggactattatt 645
Db 301 AAGCAGTTTTTCCGCAATTTTCGTGACTATTATGCTTTTGGTCTGTGTGGACTATTATT 360
Qy 646 tcttgacaatcatatctcttaggtgtaaacagcttctttaagaagttggacattggaacc 705
Db 361 TCTTGCACAATCATATCTCTAGGTGTAAACAGTCTCTTAAGAAGTTGGACATTGGAACC 420
Qy 706 tttgacttgggtgattatctctgtattgttgccatatatttctgcaacagattcagtagt 765
Db 421 TTTGACTTGGGTGATATCTTCTTATTTGGTGCCATATATTCTGTCACACAGATTCAAGTATGT 480
Qy 766 aactgcaggtctctgaataacagacgagacacaccttgtctttacagctctgttattcggagag 825
```


QY 503 atctctctgctttagtgaagatctcttcttctcatatctcttcttccaccattattca 562
 Db 911 ATCTTTTGGTCTTCAGTGAAGATCTTTTCTCATATAGCTTCTCCACCGATTATATCA 970
 QY 563 atgcaggggttcaagtaaaagaagcagcttttccgcaattctgtaactattatgttt 622
 Db 971 ATGCAGGCTTTCAGGTGAAGAAGACAGCTTCTCCGCAACTTCATTAATATGTTATGT 1030
 QY 623 ttggtgctgttggactattattcttctgacaaatcatatctctctctgtaacacagttct 682
 Db 1031 TTGAGCTGTGTGATATGATATCATATCATCATATATCTCTGGAGCGTGTGCAATTT 1090
 QY 683 ttaagaagttggacatttggacatttggacttgggtgattctctgctattgttgcattat 742
 Db 1091 TTAGAATTTGGATATTTGGTACTCTGGAGTTGGCAGACTATCTTGCATATGTCATAT 1150
 QY 743 ttgctgaacagatcttaagtatgtacactgcaggtttctgaataagaagacacatttgc 802
 Db 1151 TCGCTGCCACAGATTTCTGTTGACACTGCAGGTTCTTAATACAGATGAGACCCCTGTC 1210
 QY 803 ttacagctcttatttctgagaggggttctgtaagtgcacacgtcagttgtgctctca 862
 Db 1211 TCTACGCTGGTCTTTGGCGAGGCTGTTAATGATGCGCACATCTGCTCTTTTCA 1270
 QY 863 acggagctcagagcttggatctcaactcaactcaacacagaagctgtcttctatctcttg 922
 Db 1271 ATGCAATTCAGAGCTTTGACCTCACAGAAATTTGATCACAGAAATAGCTTTACAATTTATGG 1330
 QY 923 gaaactctgtatttcttctctactgaacttctgtgtgctgctgacacagctgtgataa 982
 Db 1331 GCAACTCTTATATTTATTCGCAAGCAGCATACTTGGAGCAATTTACTGGCTTGCTCA 1390
 QY 983 gtgagtgatttaacaagaactaacttgggaagcaactcaactcaacagagatttgc 1042
 Db 1391 GTGCTTACATTTCAAAAGCTACTTCTGGAAGGCAATTTCCACTGATCGTGAAGTTGCTT 1450
 QY 1043 ttatgactgttggcgtattcttcttcttctgctgagcttcttctgacttgagcggtta 1102
 Db 1451 TAATGATGCTTATGGCTTATCTATCTTACATGCTTGTGCAACTTTTCTTATTTAGTGAA 1510
 QY 1103 tctcaactgttcttctggttattgtatgtatctccattacacatgacgaataaagc 1162
 Db 1511 TTCCTACGTATCTTCTGAGGATGTCATGTCCTTATACCTTGGCAGCAATGTGACAG 1570
 QY 1163 agagctcaagaataacaacaaacacacttcttgcacttcttcttcttcttcttcttcttct 1222
 Db 1571 AGAGCTCAAGAGTAACCAACCAAGCATGCTTTTGCACACTATCTTTTCTGCTGAGGTT 1630
 QY 1223 ttatttctgtatgttggatgagcttcttcttcttcttcttcttcttcttcttcttcttct 1282
 Db 1631 TCCTATTTCTATATGTCGGTATGATGTCATGTCATGTCATGTCATGTCATGTCATGTCATG 1690
 QY 1283 acacacgggaacatcgatcgagtgagctcaactcaactcaactcaactcaactcaactcaact 1342
 Db 1691 ATAGTCTCGCATTTCTGTTCTGCTGAGTGTCCATATTTAGTCTAGTCTAGTCTAGTCTAG 1750
 QY 1343 aagcaggtctgcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1402
 Db 1751 GAGCAGCTTTTGTTCCTCTTATCTCTGTTGATGACCTTTCGCCAAGAAATCGCAAGTG 1810
 QY 1403 agaaaatacaactttaacacagcttcttcttcttcttcttcttcttcttcttcttcttcttct 1462
 Db 1811 AAAAGGTTCATTTCAACAGCAGATTTGTCATATGTTGGCTGGCTGCTTATGAGAGTGCTG 1870
 QY 1463 tatcttggcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1522
 Db 1871 TTTCATGCGCACTTGTCTTATTAATCAGTTTACGAGGTCTGGCAGACACACAGCTTAGG 1930
 QY 1523 atgcaatcatgacacagatcagcttcttcttcttcttcttcttcttcttcttcttcttcttct 1582
 Db 1931 ATGCAATCATGATCACAGCACTATATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1990
 QY 1583 tgcagcaaacactcataagctacctaattacccgacacagacagccacc-----ga 1636

Db 1991 TGCTGACAAAGCCTCTCATCATGTTTTCCTGCTCCCTCAACCGAAACACTTTCATGTTGCA 2050
 QY 1637 gcatgttatctgatgacaacacaccccaaat-----cc 1668
 Db 2051 GCACCGTATCAGATGTGGGAAGTCCAAAGTCATCTGTTGCCACTCTTGTGAGGCAACC 2110
 QY 1669 atacatacccttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1726
 Db 2111 AAGATTATGAAGTTGATGTGGGAACCGAAACCATGTAAGACACCACTGAGCGCGGACTA 2170
 QY 1727 tgcctcgccctgcagctatatactgcttcttcttcttcttcttcttcttcttcttcttcttct 1786
 Db 2171 TAGTTGCACTAGTAGCTCCGATGCTTCTTAATGCACTTACTACACCGCTCCATCATCT 2230
 QY 1787 actgagacaatttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1846
 Db 2231 ATTGGCCAAATTCGATGACTCTCTCATCGCGCGCTTTTGTGGTGGCGGGGTTTGTATC 2290
 QY 1847 ccttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1880
 Db 2291 CTTTGTCCCGGTTCACTTACTGTAACAAAGCAC 2324

RESULT 4

AB033989 2218 bp mRNA PLN 06-OCT-2000
 LOCUS ipomoea nil Purple mRNA for Na+/H+ exchanger, complete cds.
 DEFINITION AB033989
 ACCESSION AB033989
 VERSION AB033989.1 GI:10716128
 KEYWORDS
 SOURCE ipomoea nil (strain:KKZSK-2) fully colored flower buds cdna to mRNA, clone lib:cdna library.

ORGANISM

Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Convolvulaceae; Ipomoea.

REFERENCE

1 (sites)
 Fukada-Tanaka S., Inagaki Y., Yamaguchi T., Saito N. and Iida S.
 Colour-enhancing protein in blue petals
 Nature 407 (6804), 581 (2000)
 MEDLINE 20487008
 2 (bases 1 to 2218)

AUTHORS

Fukada-Tanaka S., Inagaki Y., Yamaguchi T. and Iida S.

TITLE

Submitted (26-OCT-1999) to the DDBJ/EMBL/GenBank databases.
 Yoshishige Inagaki, National Institute for Basic Biology, Division of Gene Expression and Regulation 1: Myodaiji-cho, Okazaki, Aichi 444-8585, Japan (E-mail:yinagaki@nibb.ac.jp, Tel:81-564-55-7684, Fax:81-564-55-7685)

JOURNAL

FEATURES

source

1. .2218
 /organism="Ipomoea nil"
 /strain="KKZSK-2"
 /db_xref="taxon:35983"
 /clone_lib="cdna library"
 /tissue_type="fully colored flower buds"
 300..1928
 /gene="Purple"
 300..1928
 /gene="Purple"
 /codon_start=1
 /product="Na+/H+ exchanger"
 /protein_id="BAB16380.1"
 /db_xref="GI:10716129"

gene

CDS

/translation="MAFGLSLLONSLLFTSDHASVVMNLFVALLCACIVLGHLLLEE
 NRWNESITALLIGCTGVILLISGCKSHLLVSEDLFLYLLPLIINAGFOVKK
 KQFVNFMTLMFLGALIGLISGSIISPGAVKIFKHLDDIFDLGDIILAGAFPAVDS
 YLPLQVLSQDETFLLYSLVFGGVDVATVFLVFNATQSFDMTSFDPKIGLHIGNFL
 YLPLSTFLVGYGLICAYIKKLYFGRHSDREVALMLMSYLSIMAEFLYSGIL
 TFPFCIVMHSYTHWNTHSRVTRHSAFTFLYVGMADALDKWFEVK
 NSOGLSVAVSSILVGLILVGRAPVFLPSLNLAKNSDKISFQQIILIMWAGLMR
 GAVSIALAVNKTTSCHTSLHENAIMITSTVTTVVLESTVTVVFLGMLTKPLINLLPLPHKQ


```

source      1...2313
/organism="Oryza sativa"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
297...1904
/gene="OsNHX1"
297...1904
CDS
/gene="OsNHX1"
/translation="similar to yeast sodium/proton exchanger"
/codon_start=1
/protein_id="BA083337.1"
/db_xref="GI:573137"
/translation="MGMEVAARLALGALYTTSDYASVVSINFLVALLCACIVLGLLEE
NRWNESITALLIGLGVVILLMTKSKSHLFFVSEDLFFYLLPPIFNAGFQVK
KOFFRNMFTITLFGAVGTMSFEFTIAAIAISFNMIGTLDVGFIAIGAFSADS
VCTQLVNMDDTPEPLSLVPEGVNDASIVLENALONFDLVIIDAAYVLKPLGNFF
YLFSSPLGVFAGLLSAYIILKLYICRHSIDREVALMLMAYLSYMLAEILLDLSGIL
YFPCGIVMSHYWHNTESRVTTHKAFATLSFAETFELEYVGMKALDIEKWEFAS
DPRKSGISILIGLVIGRAAFVPLSFLNTRKAPNEKITWRQOVVFWAGLMR
GAVSIALAYNKFTKSGHTQURNAIMITSTITVLFSTWFGMMTKPLIRLLPASGH
PVTSEPSKLSHPLSTSMQSDLESFTTNIVRPSLRMLLTTPHTVHYWRKEDDA
LMRPMFGRGFVPPSPGSPTEQSHGGR"
2313
polyA_site
BASE COUNT 543 a 538 c 569 g 563 t
ORIGIN

Query Match      35.3%; Score 768.8; DB 12; Length 2313;
Best Local Similarity 68.6%; Pred. No. 1.2e-160;
Matches 1077; Conservative 0; Mismatches 487; Indels 6; Gaps 1;

QY 312 qcctcgttcatgacatctgacagcctctgtggttgcgttgaatctcttctgtgacat 371
DB 329 GGGCTGTGTACACGACCTCCGACCTAGCGCTGGTGTGCTCATACACCTGTGCTGCGCT 388
QY 372 tcttbtgttattgtcttctgttgcacatcttcttggagaagaatagatggatgaacatc 431
DB 389 GCTCTGCGCTGCATCTGCTCGGCACCTCTCGAGGAGAAATCGCTGGTCAATGAGTC 448
QY 432 catcacgcctctgtgattggctgagcactggtgttaccattttgttggattagtagaagg 491
DB 449 CATCACCGCGCTCATCATCGGCTCTGACCGCGCTGGTGTGATCTGCTGATGACCAAGG 508
QY 492 aaaaagctcgcattctctgttctttagtgaagactattattcttgcacatactcttgcac 551
DB 509 GAAGAGCTGCGACTTATTCGTCTTCTCAGTGAGGATCTCTTCTCATCTACCTCCCTCC 568
QY 552 cattattatcaatgcagggtttcaagtaaaagaaagcagtttttccgcgaatttctgac 611
DB 569 GATCATCTTCAATGACAGGTTTTCAGGTAAGAAAAGCAATCTCTCCGGAAATTCATGAC 628
QY 612 tattatgcttttggctgttgggaactattattcttgcacatactctcctagtgct 671
DB 629 GATCACAATTATTTGGAGCGCTGGGACCAATGATATCTTTTTCACAATATCTATTTGCTG 688
QY 672 aacacagttctttaaagaagtggacattggaacctttgacttgggtgattattcttctat 731
DB 689 CATTTGCAATATTCAGCAGAATGAACATTTGGAACGCTGATGTAGGAGATTTTCTTCAAT 748
QY 732 tgggtgcataatttgcgaacagattcagatgttacactgcaggttctgaatcaagacga 791
DB 749 TGGAGGCACTTTTCTGCGACAGATTTCTGCTGACCATATTCAGGCTCCTCAATCAGGATGA 808
QY 792 gacaccttttgcattacagcttctgttattcggagaggtgtgtgaatgatgcaacgctcagt 851
DB 809 GACACCCCTTTTGTACAGTCTGTTATTCGGTGAAGGTGTTGTGAACGATGCTACATCAAT 868
QY 852 tgggtgttcaacgcatcagagcttcttgcactcactcaactcaacccagaagctcttt 911
DB 869 TGTGCTTTTCAACGGCACTACGAACCTTTGATCTTGTGTCACATAGATCGGCTCTGCTCT 928
QY 912 tcatcttcttggaaactcttctgtatttcttctcctaagtaaccttctgtgtgctgcaac 971

```

RESULT 6
 LOCUS TM021B04/c
 DEFINITION Arabidopsis thaliana BAC TM021B04.
 ACCESSION AF007271

12-JUN-1997

```

VERSION AF007271.1 GI:2191181
KEYWORDS
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
REFERENCE Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
AUTHORS Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
JOURNAL Brassicales; Brassicaceae; Arabidopsis.
TITLE 1 (bases 1 to 90019)
AUTHORS Dante, M., Wansley, P. and Gibson, A.
JOURNAL The sequence of A. thaliana TM021B04
TITLE Unpublished (1997)
AUTHORS 2 (bases 1 to 90019)
JOURNAL Washington University Genome Sequencing Center.
TITLE The A. thaliana Genome Sequencing Project
AUTHORS Unpublished (1997)
JOURNAL 3 (bases 1 to 90019)
AUTHORS Waterston, R.
JOURNAL Direct Submission
TITLE Submitted (06-JUN-1997)
JOURNAL Submitted by:
COMMENT Genome Sequencing Center
Department of Genetics, Washington University,
St. Louis, MO 63108, USA
e-mail: rwilson@watson.wustl.edu

MAPPING: Clones were assigned to the YAC map by hybridization by
M. Lohdi, Cold Spring Harbor Laboratories, and fingerprinted
by M. Marra, WashU, to pick the best candidates for sequencing.

NOTICE: This sequence may not be the entire insert of this clone.
It may be shorter because we only sequence overlapping sections
once, or longer because we provide a small overlap between
neighboring submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded or sequenced with an alternate
chemistry; an attempt was made to resolve all sequencing problems,
such as compressions and repeats; all regions were covered by
sequence from more than one subclone

NEIGHBORING COSMID INFORMATION:

The actual start of this clone is at base position 1 of TM021B04 ;
the actual end is at base position 90019 of TM021B04. The
orientation of this clone is unknown.

NOTES:
Coding sequences below are predicted from computer analysis, using
the program Genefinder(P. Green and L. Hillier, ms in preparation).
Location/Qualifiers
I..90019
/organism="Arabidopsis thaliana"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="v"
/clone="TM021B04"
/map="5"
1802..3685
/gene="A_TM021B04.11"
join(1802..2050,2468..2686,2758..2844,2901..3095,
3275..3685)
/gene="A_TM021B04.11"
/note="contains similarity to Synecococcus PCC7942
chromosomal region used as basis of neutral siteII
recombinational cloning vector (PID:g1174192)"
/codon_start=1
/evidence-not_experimental
/protein_id="AAB61082.1"

/db_xref="GI:2191197"
/translation="MWMLLHSVGLIPCSNQOKSFLFHSYRYRCIVCSSTGLSIRQ
ALEQVDSKSSGDERAALSIVKDLQKPGURCFGAARQVPRQIRYIIEELKNGINAA
SLSPDTDTTSTIERNLQIAAVGGIVAMKAFDLSQQQLFTLGFMEFLDLGVNPN
GGISLVLDITGHTFSQRYNRVQKHYYIIFHTYVCELSQHEAGFLVAYLVGLPFR
TEYLSSLEALQKESGSLNIQAGSAFYDFEFLPNKLCILFQNLNRFSCIALAGVA
TEYLLGYABGGDDISKVSFLPLKNSSDYVMLYGVFLMEQLDGLVKSGLGTOKK
ADSOVWSVLNTILLRRHEIARSKLAQAMSKGESVGSICQIIEEDSDPSDI"
complement(3906..4727)
/gene="A_TM021B04.14"
complement(join(3906..4337,4418..4497,4601..4727))
/note="coded for by A. thaliana cDNA T22079"
/codon_start=1
/protein_id="AAB61076.1"
/db_xref="GI:2191191"
/translation="MEATSLSSAATIISSSSPLSIFSPKRTDSSPPRIVRLSNKK
EDKDYPOHSESSSLFRNRYTSLNDEAMGLVLSAASKVGMWTCGMEGQSLPAKDTT
DTVSTFPWSLFTKSPRRMRVATCTNCGRTTRAINPHATDTGIVFVCGCGGVFK
LVDNLNLFHEVRYIVSSSSFDYTDARWDVSGNLDFDEDDONAGSDNDVFLP"
complement(12650..16625)
/gene="A_TM021B04.10"
complement(join(12650..13129,13415..13509,14187..14498,
14596..16014,16127..16625))
/gene="A_TM021B04.10"
/note="similar to N. tabacum membrane-associated
salt-inducible protein (PID:g473874)"
/codon_start=1
/evidence-not_experimental
/protein_id="AAB61077.1"
/db_xref="GI:2191192"
/translation="MKSDFLTSTHTFNFSLFIPKIPSRNSRISIKSSSSSKVRPDPW
SLSDGNPEKPKPRYERPKHPLDODDARILIKKAOYLSTLRNOCQSOAMTPKWKTP
EQMWYKLEDDNRNGMYGKHVAAIKTVRGLSORROGSDDMREVMSSVAKLSRDMCV
VLKEQGEFCVKQLSVRPSVVYTVLRLYGWGGIKMAEETFLMEVLECEGPAVACG
TMCTVARGRHSAMLTFFKAVQERILLSTSYNFMSSLSQKSKFGICKVIDLWLEWY
EGVPPNPFYTLVSSYAKGQKEALKAFGEMKSLGVFPEEVTIISVLSLVKAGD
WEKAIYEDMRSGQIVPSNTCATMLSYIKTYNTKALSFADMRNKNIPADEVIR
GLIRYKLGLEDAQSMFEETRLNLIADEKTYLAMSQVHLNSGNVKAIDVEMM
KTRDIPLSFAYTVMQLQYAKIQNDCAEFALSKTGLPDASSNDMLNLYRNL
GEKAGFIKQIMDVQHFDTIELYKTAMRYCKEGVAEODLIVMGREARVDNRFL
QTLAESMHVINKHDKHEAVNSQLDMVALGMLNLRLKGNLNETKAILNLKFKDL
GSSAVNRVISVFRGDDVSKAEMADIIRLGLRMEETIATLIAYVGRQHLKEAR
LYLAAGESKTPGKSVIRMDAYVRCGLMEDAYGLFESAPKGCDCAVTISILVNL
TNELGLYDEKITYNMIMHYGKGMYLIKFTNMLRPSYNMVKICATSRLLHVEDL
QAMERNGRCDLSTYTLILQVIAESSQFAEAKITILVKEGIPLSHSHSSLSALY
KAGMEAEKTYCKMSEAGISPDSSACKRTILKGYMTGDAEKGILFYERKMIRSVEDD
RFVSSVVEDLYKAVGKEQDV"
complement(17345..24723)
/gene="A_TM021B04.9"
complement(join(17345..17911,17988..18340,19797..20049,
20117..20610,21561..24723))
/gene="A_TM021B04.9"
/note="contains similarity to a DNAJ-like domain"
/codon_start=1
/evidence-not_experimental
/protein_id="AAB61072.1"
/db_xref="GI:2191187"
/translation="MDWNKEACRAKTLAEDKMEGDFVGAQKLLKQAQSLFSLBSL
POMLAVCDVHNSAEKINCLENWYGILOVMHFADDAIKKORLKLALLHPDKNQFPG
AEAKLWVDSARFLADKRQYDITRRIRYLRLATNOLNANSGLQCAATNSATDTFW
TCCHECGYRYKYLRYKYNILLNCNICQSRYSMAYDTGFNEAPSKNTQKQVQNGPCN
TSLNAGESIQAQPGSVAAEVDKGTFFNKKFNKNGCGESKKTETVKNKSLTEFVRND
DEKMSAKLHAPHPFEVTEPEIGASKSVPSDESRSDEAPSTSKDNKKMKGVESI
NVGDSDFIKDAKYSKDRKSKSPQSQSSYAAEEKISDNLGPPKRLNKNVCLKSE
QTTKRLGVGSGSKDFSGSSAFSCAFKRAKKFVDSGOYEIISVSKKREVRKAS
GEGVWEAKIDNNHNNENLITEDLPDPFNFELTSCFGVQWQVSNYDPIQMPRL
YARDKVLVPEPKLWITWIDPLQNDKNSIPACIFQGGSEENHLKFSQMPRL
TRNSVYVYPRKEIWAIFRQWISWASSENHKNHPYEDFVLSNENGLGVGF
LGKVEGYSIFRODAQDGVLOLQIPPSQMLRSHKVPSPKFKRGEPVPGCGELDP
AALPKELFVYDSDVDGLDRELPGKGTGGPPEAKVEMOANTTRESSPKKRPKY
HDHGSFESKESDORRTNHNRSASKRSKRVAYDALNLKSPRLSETNSQANSRGR
QIAEKSNANSGSCGPGSLSGENKSEKTIITRMKTPQDIHAKPTVYVKKGRNDES

```

LSQSRDGLLTQINGNMKSSPEPVRPSSCKTVEKNTFENFENQSRWSDKFQIDQIWAII
SNDKSPKRYAQIKKIDTSPRFKLHVAPELVRPDIHMRPVPCGGRFLKTKAEVVV
PSFSHOWKAVKIKKIDTSPRFKLHVAPELVRPDIHMRPVPCGGRFLKTKAEVVV
QAMTITAKGFNNKPLRYRSESNASFIMERNLCBEDSDSDVIMLLMVGVSDATYESQI
PRMRKQITKGGHATIQALKDADDPHFQRLYRMMPEFAEJCHLLQMKTLGUKTGHVHC
VEMVATFELITVQNSRYCHYDFTFKRSKFTSINFHKLRLALNMLPTLMKAYTNV
PSIKSTTFYPFYKDCVGAIDGTHINAMVQPEKASYNRKRGVISOVLAAACNFDE
FYIIVLSGWEFSADHSKVLODALTRTNRLOVPEGNRLQVLKGIWMDSPKRRKGDY
NPMSPETKLLVOLLVEGINNNWRDSNCTISKLTVEKFMPEINKECRCKSNYHLS
RMKYLKIQOCSLDLORPSSGFGDPDLTKRETASDEWSDYLKHPNNKQLRYDTFFR
FDELQITIFEGVATKNAIGLCSDTGLTYRAGENPRKEYVDDFEDNVEYDITTHES
SHYAPFFWSDTESPPLPPRRKTSERSTOKESPMVSSKILLDIIQOREEQON
EYAAKNNVMDAIEKISDLDECIRYKALTKIYHLGIQDVFVYSMSVEERLGIQIME"
complement(34437. .37759)
/gene="A.TM021B04.8"
complement(join(34437. 35003,35094. 35272,35446. .36036,
36081. .36197,36313. .36607,36724. .36905,37102. .37759))
/gene="A.TM021B04.8"
/note="contain similarity to type 1 inositol
1,4,5-triphosphate receptors"
/codon_start=1
/evidence=not_experimental
/protein_id="AAB61078.1"
/db_xref="GI:2191193"
/translation="MKAKQSEFEKKEKDFLEQKAEVEKRRKEVEQLEKFTTRMESVE
RVSDKLMELGURATELEKMEVEKHEKRIYVAGDKLGRFEPLVSLKKNGLSVTM
PVKCSLVLINENADEMKNTALARMVPYLDPAKVLLDAIEGSEKYEYWKDLGEADDR
VYNSIVLLENLIKMLKITPQVQOEATPLGIWLGKAKANKNDPPVOVFCALEFLAA
YGAATLKKEBILATLAFICEFLUKICPGRGPELLIEFFDSKDAARYIAGTGS
MBAQARREKKADADAIVYKEAKETMPAKILKRLAVKNDSEARAMEPVQKS
YEKQSTTGVKEKSEAKSIPYEQKHVKRPRLTEPTAPSONLTVCFNMKHELTSMV
NLSIFVGETLRHKNONTFLFLFKILKQPEVVCPTQVKVESGADHOPDTIATHPSG
TETKLNILSGSTFKADMLRELVEKOPKLESDLSNALKCTPDPAKFLDTSWALCPTNT
EGGYEFKMLISGSSCLLNQLKLLPKLHPVKGDAKKLAVYWKDKIAKSKROLEV
ICFLQFLGIFGVSCEKADLLDLNDSYQWTVSPDICFLGLDDAIPGFIQNLKIGT
HRIKAIDIYSGMVHRFPQSAIINDSLRTKESAEKSYREAKNESTQVAADIROV
RALRAAIKICISCHLESPFQLEDQEQIKSLKLLKRLNRTSNGSGSSASSRKPDSIKOS
QTAKPPTVAEAPVTSNIPLEPSTEASASSAKPFSKKNKGRKSRMSGNOSGHIA
SHTSNHYPSHSDYSLNQLRTWPVDNYDRGTFGFPNPDYNNQWQGPQGFYHLQPLD
PRYRN"
complement(39523. .43518)
/gene="A.TM021B04.7"
complement(join(39523. .39921,40013. .40173,40283. .40858,
41109. .43518))
/gene="A.TM021B04.7"
/note="contains similarity to tropomyosin and kinesin"
/codon_start=1
/evidence=not_experimental
/protein_id="AAB61071.1"
/db_xref="GI:2191186"
/translation="MEEIKLENIIRLCDVKANNIRKTMWIKSQASDVLILNLQWDF

Query Match 31.8%; Score 693; DB 15; Length 90019;
Best Local Similarity 69.1%; Pred. No. 1.le-143;
Matches 1379; Conservative 0; Mismatches 0; Indels 616; Gaps 7;
QY 773 aggttctgaatcaagcagacaccttgcctttacagctcttgccttcgagagaggtgttg 832
Db 66754 AGGTTCTGAATCAAGACGAGACACCTTTCCTTTACAGTCTTGTATTCGGAGGGGTGTG 66695
QY 833 tgaatgatgcacagctgtgtgtcttcaacgcgattcagagcttgccttcactcaccc 892
Db 66694 TGAATGATCAACGTCAGTTGTGTCCTTCAACGCGATTTCAGAGCTTGTATCTCACTCACC 66635
QY 893 taaccacgaagctggttttcattctcttcttggaactcttgccttgccttcactaagta 952
Db 66634 TAAACCACGAAGCTGCTTTTCATCTCTTGGAACCTCTCTGTATTTGTTCTCTCCTAAGTA 66575
QY 953 ccttgccttggtgctgca----- 969
Db 66574 CCTTGTGCTGCTGCAGTAAGTCCATTTGTTCTTCCTGCCATTCCTTCTTGTAGCTG 66515
QY 970 -----accggtcgtgataag 983

Db 66514 CTATATGCACGCTCTGTGGTGTTCACATCTTTTGTATTTCATCTTGCAGACCGGTCTGATAAG 66455
QY 984 tgcgtatgttalcagaagactatcctttgga----- 1014
Db 66454 TGCCTATGTTTATCAAGAAGCTATACCTTTTGGGAAGGTACATATTAACGTCATATATATGAC 66395
QY 1015 ----- 1014
Db 66394 TTGTTTCTTTAAGTTGTTTCTCTGTGGTCTAACGTAGATGGCCTCACATTAATTTGTCA 66335
QY 1015 -aggcaactcaactgaccgagagaggttgccttctatgatccttatggcgtatctcttcttatat 1073
Db 66334 CAGGCACCTCAACTGACCGAGAGGTTGCCTTATGATGCTTATGCGGTATCTTCTCTTATAT 66275
QY 1074 gcttgctg----- 1081
Db 66274 GCTTGTCTGAGGTTAGATTCTTTTAAATAACACGAACATATTCATTCAGTTGCATTCAAAC 66215
QY 1082 -----agcttttcgacttgagcgggtatcctc 1107
Db 66214 ACAAAATATATGATGTTATATTAATTTGGTGTGAGCTTTTCGACTTTCGACGGGTATCCTC 66155
QY 1108 actgtgttttctgtggttattgtgatgctccattacacatggcacaatgtaacggagagc 1167
Db 66154 ACTGTGTTTTCTGTGGTATTGTGATGTCCTTACACATGCGCAATGTAACGAGAGC 66095
QY 1168 tcaagaatacaacaacaa----- 1184
Db 66094 TCAAGAATAACAACAAGTATGCAAACTTTAGTTTTCTCCGCTTTTGTGCGAGATTAACAC 66035
QY 1185 -----gcatacctttgcaa 1198
Db 66034 TTTTACCTCCCATGATTATCTTATGATTTTTCGATGAACAGGCATACCTTTGCAT 65975
QY 1199 ctttgtcattcttcgagagacatttatttcttctgtatgttggaatggatgccttcgtgaca 1258
Db 65974 CTTTGTCTATCTTTCGCGGAGACATTTATTTCTTGTATGTTGGAATGATGATGCGTTGGACA 65915
QY 1259 ttgacaagtggagatccgtgagtgacacacacgggaaacatcgatcgagtgagctcaatcc 1318
Db 65914 TTGCAAGTGGAGATTCCTGTGAGTGACACACCGGGAACATCGATCGAGTGCAGTCAATCC 65855
QY 1319 taatgggtctgctgctgagtgagagagcagcgttcgtctcttccttcgttctctacta 1378
Db 65854 TAATGGGTCTGCTCATGTTTGGAAAGCAGCGCTTCGCTTTCCGTTATTCGTTTCTACTA 65795
QY 1379 acttagccaagaagaatcaaaagcgagaaaaatcaactttaacatgc----- 1423
Db 65794 ACTTAGCCAAGAAGAAATCAAAAGCGAGAAAAATCAACTTTAACATGCAGTATGTATCATTC 65735
QY 1424 ----- 1423
Db 65734 TACTTGTACTAATAGATTCATAATTTCTTCTGGATTTTCAGCTATTTATATCTATCATTT 65675
QY 1424 -----aggttgatgttgcgtcgtcgtcctcatgagagtgctcgtatcatg 1470
Db 65674 TTTCCCATTTCTTAGGTTGTGATTTGGTGTCTGCTCTCATGAGAGGTGCTGTATCATG 65615
QY 1471 gctcttgatacaaca----- 1486
Db 65614 GCTCTTGCATACACAAGGTATATAAAGAAATGCTTTAAATACTCTCTATGAAGCTTCAA 65555
QY 1487 -----agttacaagggcgaggacaca 1509
Db 65554 ATGACAAAATCATCATTTAAACACTTGTCTTCTATTGTAGTTTACAAAGGGCGGACACA 65495
QY 1510 gatgtacgaggaatgaatcatcatgatcacgagtaacatacgtctgctctcttttttagaca 1569
Db 65494 GATGTACGCGGGAATGCAATCATGATCACGAGTACGATACGTGTCTGTCTTTTGTAGCACA 65435
QY 1570 gt----- 1571
||


```

TITLE      Na+/myo-inositol symporters and Na+/H+-antiporters in plant sodium
            transport
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 459)
AUTHORS    Chauhan,S., Forsthoefel,N. and Bohnert,H.J.
TITLE      Direct Submission
JOURNAL    Submitted (19-JUN-2000) Biochemistry, University of Arizona, 1041
            E. Lowell St, Tucson, AZ 85721, USA
FEATURES   Location/Qualifiers
            1..459
            /organism="Wesembryanthemum crystallinum"
            /db_xref="taxon:3544"
            <1..>459
            /gene="NRX1"
            <1..>459
            /gene="NRX1"
            /codon_start=1
            /product="putative Na+/H+-antiporter"
            /protein_id="AAF91423.1"
            /db_xref="GI:9652168"
            /translation="pPIIFNAGFQVKKQFFKNFSFILMFVGVLGTIISGLISSGVAL
            LKKGITGLNLDNLALCAIFASDTSVCTQLVNQDETPLLAYSIVFEGGVNDATSI
            VLENVAVOSDLSNLSMTALALLGFLVLYFTSTILGILGLASAYVKKL"
BASE COUNT 115 a      98 c      89 g      157 t

```

	Query Match	9.3%	Score 203.6;	DB 13;	Length 459;
	Best Local Similarity	65.3%;	Pred. No. 4.3e-35;		
	Matches 299;	Conservative 0;	Mismatches 159;	Indels 0;	Gaps 0;
Qy	547	ccaccctatattcaatgcagcagggtttccaagtaaaaaaagaacgagttttccgcgaatttc	606		
Db	1	CTCCCAATAACTTCAACGCGGGGTTTCAGGTTAAGAAAAGCAGTTTTTCAAGAAGCTTC	60		
Qy	607	gtgacattatcgctttttgtgtgctgttgygactattatttttgcacaaatcaatctctcta	666		
Db	61	TCATTCATATATTGATGTTTGGCGTACTCGGAACAATAATATCATTTTGGCCTCATATCATCA	120		
Qy	667	gggttaacacacagttctttaagaagttggacatgggaacctttgaccttgggtgattatctt	726		
Db	121	GGTGCGGTACTACTGTTGAAGAAGATTGGTTTGACTCAGCTTAATCTTAATGACTATCTG	180		
Qy	727	gctattgtgcataatttgcataagattgcacaagattcagtagtgacactgcaggttctgaatcaa	786		
Db	181	GGCCTTGGTGGCATATATCTCAGCCACAGACTCTGTTGACTTTTGCAGGTCTCTCAATCAG	240		
Qy	787	gacgagacacctttgctttacagtccttgatttcggagaggggtgtgtgaatgatgacaagc	846		
Db	241	GATGAACACCCGCTCTCTTACAGTATTGTATTGTGGAGGGGAGTAGTAATGATCAACCC	300		
Qy	847	tcagttgtggctctcaacgcgattcagagctttgatctcactcaactaaaccacgaagct	906		
Db	301	TCCATTGTGCTTTTCAATGTCAGTCCATCACTTGATTTAAGCAACCTCAGTTCATGACT	360		
Qy	907	gcttttcattcttttggaaaccttttggattttgtttctctcaagtagaccttgggtgct	966		
Db	361	GCTTTAGCATTTGCTGGGAACCTTCTCTATCTCTTCTTCCACGAGCACAAATCTTGGCATT	420		
Qy	967	gcaaccggtctcgataagtgcdtatgtattatcaagaagct	1004		
Db	421	CTAGTTGGTTTGTGGAGTGCTTATGTATTATAAGAAGCT	458		
RESULT	10				
AF307944					
LOCUS	AF307944	320 bp	mRNA	PLN	02-APR-2001
DEFINITION	Zea mays	Na+/H+ antiporter	mRNA,	partial cds.	
ACCESSION	AF307944				
VERSION	AF307944.1	GI:13507162			
KEYWORDS					
SOURCE	Zea mays.				
ORGANISM	Zea mays				

RESULT 12
 ATAC011623
 LOCUS
 DEFINITION ATAC011623 86022 bp DNA PLN 24-JAN-2001
 Arabidopsis thaliana chromosome III BAC F24P17 genomic sequence,
 complete sequence.
 AC011623
 VERSION AC011623.5 GI:12408729
 KEYWORDS
 SOURCE HTG.
 ORGANISM
 Arabidopsis thaliana
 thale cress.
 Arabidopsis thaliana
 Magnoliophyta; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 Magnoliopsida; eudicotyledons; core eudicots; Rosidae; eurosids II;
 Brassicales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 86022)
 Lin.X., Kaul.S., Town.C.D., Benito.M., Creasy.T.H., Haas.B.,
 Ronning.C.M., Koo.H., Fujii.C.Y., Utterback.T.R., Barnstead.M.E.,
 Bowman.C.L., White.O., Nierman.W.C. and Fraser.C.M.
 Arabidopsis thaliana chromosome III BAC F24P17 genomic sequence
 Unpublished
 2 (bases 1 to 86022)
 Lin.X. and Kaul.S.
 Direct Submission
 Submitted (08-OCT-1999) The Institute for Genomic Research, 9712
 Medical Center Dr., Rockville, MD 20850, USA, xlin@tigr.org
 3 (bases 1 to 86022)
 Lin.X.
 Direct Submission
 Submitted (24-JAN-2001) The Institute for Genomic Research, 9712
 Medical Center Dr., Rockville, MD 20850, USA
 On Jan 24, 2001 this sequence version replaced gi:12280803.
 Address all correspondence to:
 Xiaoying Lin
 The Institute for Genomic Research
 9712 Medical Center Dr.
 Rockville, MD 20850, USA
 e-mail: xlin@tigr.org
 BAC clone F24P17 is from Arabidopsis chromosome III and is near the
 molecular marker mi403.
 The orientation of the sequence is from SP6 to T7 end of the BAC
 clone.
 Genes were identified by a combination of three methods: Gene
 prediction programs including GRAIL (available by anonymous ftp
 from arthur.ornl.gov), Genefinder (Phil Green, University of
 Washington), Genscan (Chris Burge, and NetPlantGene
 http://www.cds.dtu.dk/netgene/cbsnetgene.html), searches of the
 complete sequence against a peptide database and the Arabidopsis
 EST database at TIGR (http://www.tigr.org/tdb/at.html).
 Annotated genes are named to indicate the level of evidence for
 their annotation. Genes with similarity to other proteins are named
 after the database hits. Genes without significant peptide
 similarity but with EST similarity are named as 'unknown' proteins.
 Genes without protein or EST similarity, that are predicted by more
 than two gene prediction programs over most of their length are
 annotated as 'hypothetical' proteins. Genes encoding tRNAs are
 predicted by tRNAscan-SE (Sean Eddy,
 http://genome.wustl.edu/eddy/tRNAscan-SE/). Simple repeats are
 identified by RepeatMasker (Arian Smit,
 http://ftp.genome.washington.edu/RM/RepeatMasker.html). Regions of
 genomic sequence that are not annotated as genes but have predicted
 exons by GRAIL are annotated as misc features.

FEATURES
 Location/Qualifiers
 1..86022
 /organism="Arabidopsis thaliana"
 /cultivar="Columbia"
 /db_xref="taxon:3702"
 /chromosome="III"
 /map="mi403"
 /clone="F24P17"
 join(<283..529,638..818,916..1063,1101..1142,1160..1376,
 1511..1822)

repeat_region
 5843..5962
 /rpt_family="(TGG)n"
 complement(7077..>8090)
 /gene="F24P17.3"
 complement(7077..>8090)
 /gene="F24P17.3"
 /note="similar to hypothetical proteins GB:CAB16818,
 GB:AAD5277"
 complement(7254..8090)

mRNA
 join(<283..529,638..818,916..1063,1101..1142,1160..1376,
 1511..1822)

gene
 /gene="F24P17.1"
 <283..1822
 /gene="F24P17.1"
 /note="similar to pyruvate dehydrogenase kinase
 GB:AAC97601, [Arabidopsis thaliana]"
 CDS
 join(<283..529,638..818,916..1063,1101..1142,1160..1376,
 1511..1569)
 /gene="F24P17.1"
 /codon_start=1
 /product="putative pyruvate dehydrogenase kinase, 5'
 partial"
 /protein_id="AAF08568.1"
 /db_xref="GI:6437541"
 /translation="VRDYLSEPRMRAPPEIKDSGDEKDFOTMIKAVKVRHNNVPM
 MALGVNOLKKGMSGNDLDFYLSRIGRIGRLMGHQLHVNPNPLHTVGYTH
 TKMSPMVARNASDARSICFREYSGAPEINIGDPSTFFYVPTHLHLMYELVKS
 LRAOERFVDSRVAPIRTIIVADGLEDTIKPERSLLRREDPIIVVSDEGGIARSG
 LPRIFTLYSTARNPLEEDVDGIADVPVTMAGYCYGLPISRLYARVFGDGLQIISME
 GYTDALHLSLRGDSQEPUP"
 complement(2022..2051)
 /rpt_family="POLY_A"
 complement(2053..2372)
 /note="exon predicted by xgrail, quality
 excellent_shadowexon"
 repeat_region
 2057..2160
 /note="exon predicted by xgrail, quality excellent"
 misc_feature
 2287..2363
 /note="exon predicted by xgrail, quality excellent"
 misc_feature
 join(<2630..3712,3813..4007,4256..4324,4406..4675,
 4767..5117,5215..5492,5579..>6599)
 /gene="F24P17.2"
 <2630..>6599
 /gene="F24P17.2"
 /note="similar to RNA helicase (DRH1) GB:BAA28347
 [Arabidopsis thaliana]; Pfam HMM hits: WW/rsp5/WWP domain
 containing proteins, DEAD and DEAH box helicases,
 Helicases conserved C-terminal domain"
 CDS
 join(2630..3712,3813..4007,4256..4324,4406..4675,
 4767..5117,5215..5492,5579..>6599)
 /gene="F24P17.2"
 /codon_start=1
 /product="putative RNA helicase"
 /protein_id="AAF08584.1"
 /db_xref="GI:6437557"
 /translation="MATTETPASAGPRVAPEDPTLPQPKGLIDSGTGLIYWNPE
 NVTQYRPSAPPHSATTPKLAOIYPVSSGOGHQHQAKPVGHVSQHQGFQOQPOO
 PPSVHPQMQHQHPAQMPQSQSQFPOQSQSMVPHGHPSTVQTYQTPTTQQOQOQ
 MNQHSMPQQLSHQYAHSQHQHGFPHMQTGLQNSHQTPQCGHGOFPQSQOQY
 SLAPRDEDFHGGKTKFSQPHLPSERSPSQNTHEANASQSTNNANLAMAQKNG
 POANAAVTQOQGANLIHQOLGPRAPNQMDQTLHOKSHVSPQNNYENNLQSRP
 GNDSYVMRVEVPVGAQPLHPAAMPKDJIRISGGPTNADPAMQGTGHGTGHAGAPR
 PKSLVPRHFTVSPVPHLSPVEIYRKOHEVTTGTENIPAPYITPESGLPELREL
 LSAGFPSPPTIAQATWPIALQSRDIAIAITGSGTKLYLIPAFILLRHRNDRNGP
 TVLLAPTRLATIODEALRCRSRISCTCLYGAPKPOLKELEGADIVATWPG
 RLNDLEMKIDTFOVSLVLDRAHMLDMFEPOIRKIVNEIPPRPTLMYATWPK
 EVRTASDLLVNPQVNIQVDELAANKALTOYVEVPMKERLEOILRSQRGSK
 VIIFCSKRLCDHLARSYGRHGAIVYHGDKTQGRDVAQLNGFSGKSCVLATDVA
 KGLDIKOLRVINIDFTPTGVEDYVIRHIGTGRAGATGVAFTEFTEDQWKYAPDLIKVL
 EGAKQVPPQVDRIAMRGGGGGPGYSQRRRNKVFUSGGGTQWDSGGGFGGGGG
 FSGREGFGGREGGFGGREGGFGGRCMDRDSFGGNGRGRGTGPDAGHNNVGG
 RGGCRGNNNNMESRGFGGSGRGFGGRCMDRDSFGGNGRGRGTGPDAGHNNVGG
 RSRSGSYSRSRSRSRSRSRSRHRSRDRGRGHNRSYSYRSRSPVYRRRRAPR
 VSGFDIKPPVSVVNLDMNAAAIAENNVPTLSERGNGVVESEAAALVPVYVDEEP
 " 5843..5962
 /rpt_family="(TGG)n"
 complement(7077..>8090)
 /gene="F24P17.3"
 complement(7077..>8090)
 /gene="F24P17.3"
 /note="similar to hypothetical proteins GB:CAB16818,
 GB:AAD5277"
 complement(7254..8090)

ACCESSION U33007.1 GI:927685
 VERSION 1
 KEYWORDS Saccharomyces cerevisiae strain-S288C (AB972).
 SOURCE Saccharomyces cerevisiae
 ORGANISM Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 REFERENCE 1 (bases 1 to 78500)
 AUTHORS Dietrich, F.S.
 TITLE The sequence of *S. cerevisiae* lambda 3641 and cosmids 9461, 9831, and 9410
 JOURNAL Unpublished (1995)
 REFERENCE 2 (bases 1 to 78500)
 AUTHORS Dietrich, F.S., Mulligan, J., Allen, E., Araujo, R., Aviles, E., Berrio, A., Carpenter, J., Chen, E., Cherry, J.M., Chung, E., Duncan, M., Hunnicke-Smith, S., Hyman, R., Komp, C., Lashkari, D., Lew, H., Lin, D., Mosedale, D., Nakahara, K., Namath, A., Oefner, P., Oh, C., Petel, F.X., Roberts, D., Schramm, S., Schroeder, M., Shogren, T., Shroff, N., Winant, A., Yelton, M., Botstein, D., and Davis, R.W.
 TITLE The sequence of *Saccharomyces cerevisiae* chromosome IV right
 JOURNAL Unpublished (1995)
 REFERENCE 3 (bases 1 to 78500)
 AUTHORS Dietrich, F.S.
 TITLE Direct Submission
 JOURNAL Submitted (02-AUG-1995)
 REFERENCE 4 (bases 1 to 78500)
 AUTHORS Jia, Y. and Cherry, J.M.
 TITLE Direct Submission
 JOURNAL Submitted (12-JUN-1997) Department of Genetics, Stanford University, Saccharomyces Genome Database, Stanford, CA 94305-5120, USA
 COMMENT Sequenced by: Stanford DNA Sequence & Technology Center
 855 California Avenue
 Palo Alto, CA 94304, USA
 Curated by: Saccharomyces Genome Database
 URL: <http://genome-www.stanford.edu/>
 e-mail: yeast-curator@genome.stanford.edu

Neighboring Sequence:
 The 5' end of this sequence contains a 850 bp overlap with GenBank Accession Number U32274.

FEATURES
 source
 1..78500
 /organism="Saccharomyces cerevisiae"
 /strain="S288C (AB972)"
 /db_xref="taxon:4932"
 /chromosome="IV"
 /map="4P"
 /strain="S288C (AB972)"
 1..42652
 /organism="Saccharomyces cerevisiae"
 /db_xref="taxon:4932"
 /clone="cosmid 9461"
 <1..11552
 /organism="Saccharomyces cerevisiae"
 /db_xref="taxon:4932"
 /clone="lambda 3641"
 complement(603..1322)
 /gene="STE14"
 complement(603..1322)
 /gene="STE14"
 /note="STE14"
 /codon_start=1
 /product="Ste14p: farnesyl cysteine carboxyl-methyltransferase"
 /protein_id="AAB64880.1"
 /db_xref="GI:927714"
 /translation="MHODFODEHEYPDIRRNPLHEWTMTVSILGILLGIFGLFPQI
 RFKNFLNFIALLFHTLEYITAKYNPLKVHSEFLNNGKSYNAHSEALLLECLVE
 SFLFDPLKIFSYSLATKLCIVLCGLVILGQYTRTAMTAGHSFHVYTKKESDHV

LVKTGVYSGRHPSYLGFFWMAIGTQLLLNPLSLVIFVLMKFFSDRIRVEKYLI
 EFFSAEYIEYKNKVGVPPI"
 complement(1600..2625)
 /gene="YDR411C"
 complement(1600..2625)
 /gene="YDR411C"
 /note="similar to Caenorhabditis elegans hypothetical
 protein from GenBank Accession Number U00036, CDS number
 7; CAI: 0.10"
 /codon_start=1
 /evidence=not_experimental
 /product="Ydr411cp"
 /protein_id="AAB64889.1"
 /db_xref="GI:927723"
 /translation="MAGRNVRVLHNGGRNNDVMGPKFVWLNIPTRTFLTLAIVM
 TIVGRNLINPWFYIYVNTFKKYLWLLTSCVMSRRAMPALMEIYSIDRSSOL
 ERGFGGLGNRRGPMVVDYAYLFCILAITATTIYGSYVPVYSIFISCIY
 TWSIDNANQIMFYGLIPVNGKYPFLQIFGIFVNEGDFVLSLIGFTGLYLCLOT
 HTLGPIMGMSRKADPTYGISPNKFTSPWMTSLYARITGAHNETAFNNFANVS
 SORTRVFSGRGQLGTAPATLSQTSITDGRASGSQLRSGLNQFQGRQVGGT
 NSPDSQ"
 2924..3631
 /gene="YDR412W"
 2924..3631
 /gene="YDR412W"
 /note="CAI: 0.16"
 /codon_start=1
 /evidence=not_experimental
 /product="Ydr412wp"
 /protein_id="AAB64855.1"
 /db_xref="GI:927689"
 /translation="MAVNTNRQILTRKNTATQSKKFGTDEVTFDKSRDLYLTGFH
 KRKLQKKAQEFKQEERLKIIEERQIRQERKEVMEQLTKFESLNLEAIEDAK
 NDKTEDLQVDESDESHGFDSDKDDGNDNNSSVVKPLKKGAITIYDSTTVELET
 EPNDNFYLAQLNNVKLEAKYLVKQISINRAIKYAKFLGVDEKOKKPRVKRYLTK
 NERRINORKANDNKKRR"
 complement(3067..3642)
 /gene="YDR413C"
 complement(3067..3642)
 /gene="YDR413C"
 /codon_start=1
 /product="Ydr413cp"
 /protein_id="AAB64892.1"
 /db_xref="GI:2194165"
 /translation="MRIHYLLRLSPAFILRSLFKVLYLNFFTLGFFGFCFSSTPRN
 FAYLVAFMLCLSTFSFSLNLTFSWAKYSKLSGNSVNSSTVSESSYISVIAPEFKI
 GFTLISLISLSPSSLSSESNPCQSDSTCKSVLSFSPASSISASAKFLSLNWFNCSSI
 TSLRSCRIFCLSSITFLSLSCSLNSCAFFCL"
 complement(3829..4917)
 /gene="ERD1"
 complement(3829..4917)
 /gene="ERD1"
 /note="required from retention of luminal endoplasmic
 reticulum proteins; YDR414C; CAI: 0.11"
 /codon_start=1
 /product="Erd1p"
 /protein_id="AAB64888.1"
 /db_xref="GI:927722"
 /translation="MEKSESNSEGLYQLNLNPPQRFIVLIILALMTWLKFFL
 HSNLDVSQVILTRVPHDIRGYTLQQLHRTARNFALKITRIIPFHAIVLEFFEMNI
 IEGPLKNILIVLEPLIQCVTFWLLKECQIKYCTRCLLIESSPRLNRYILI
 SDTLTSGKPLIDFTLSTLIFREPFTFDFLSVALLPLVLRLLQCLREYRLHEATLL
 FNALKYSCNLPILFCTWRSRVSINEERLHHVQRFMLINSYTLFWDVMDWSLD
 SLTSLRGRSKSAVTLKKKMYHSAILVDFLLRFWMLVWVLSQNLKLVAAADSDYIFFQGE
 MOYFEVIRGIWVVFKLDAEYIKFASK"
 complement(5268..6392)
 /gene="YDR415C"
 complement(5268..6392)
 /gene="YDR415C"
 /note="similar to bacterial leucyl aminopeptidase from PIR
 Accession Number A38086; CAI: 0.14"
 /codon_start=1
 /evidence=not_experimental

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2001, 13:26:37 ; Search time 22.15 Seconds
(without alignments)
1472.492 Million cell updates/sec

Title: US-09-271-584A-2

Perfect score: 2755
Sequence: 1 MLDSLVKPLSLSTSDHASV.....FVPFVPGSPTEPNPPDLKA 538

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 10% summaries

Database :

A_Geneseq_0601.*
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	DB ID	Description
1	2755	100.0	538	20	AA19901	Arabidopsis thaliana
2	2516	91.3	547	20	AA19905	Arabidopsis thaliana
3	2168.5	78.7	555	21	AA12786	At11px gmelini N
4	2091.5	75.9	553	22	AA173253	Protein regulating
5	2059.5	74.8	540	22	AA173252	Protein regulating
6	2041	74.1	542	22	AA173251	Protein regulating
7	2023.5	73.4	535	21	AA11832	Rice Na+/H+ antipo
8	1982.5	72.0	555	22	AA173254	Protein regulating
9	1570	57.0	529	20	AA19902	Arabidopsis thaliana
10	582	21.1	608	21	AA129621	Cat flea HMT Na/H
11	565	20.5	526	22	AA190637	Human secreted pro

12	564	20.5	633	20	AA144017	Saccharomyces cere
13	563	20.4	509	22	AA190555	Human secreted pro
14	544	19.7	569	20	AA144016	Schizosaccharomyce
15	463	16.8	834	18	AA111325	Human Na+/H+ excha
16	441	16.0	816	14	AA144143	Rabbit sodium ion/
17	429	15.6	325	20	AA144096	Arabidopsis thalia
18	320	11.6	339	22	AA190590	Human secreted pro
19	320	11.6	339	22	AA190591	Human secreted pro
20	263	9.5	232	20	AA140907	Arabidopsis thalia
21	237	8.6	185	20	AA140903	Arabidopsis thalia
22	151	5.5	79	21	AA144993	Human secreted pro
23	144	5.2	511	22	AA176843	Corynebacterium g1
24	143.5	5.2	178	20	AA136089	Extended human sec
25	139	5.0	178	20	AA135990	Extended human sec
26	137	5.0	480	19	AA198431	H. pylori GRPO 446
27	136	4.9	129	21	AA153670	Human colon cancer
28	129	4.7	412	21	AA170156	Staphylococcus aur
29	119	4.3	443	21	AA144094	Arabidopsis thalia
30	119	4.3	518	21	AA144093	Arabidopsis thalia
31	119	4.3	568	21	AA144092	Arabidopsis thalia
32	119	4.3	697	21	AA193290	Amino acid sequenc
33	117.5	4.3	430	20	AA188344	Salmonella enteric
34	117.5	4.3	469	18	AA189797	Staphylococcus aur
35	117.5	4.3	482	13	AA121409	NADH dehydrogenase
36	117	4.2	440	21	AA133314	Arabidopsis thalia
37	117	4.2	440	21	AA146728	Arabidopsis thalia
38	117	4.2	515	21	AA133313	Arabidopsis thalia
39	117	4.2	515	21	AA146727	Arabidopsis thalia
40	117	4.2	565	21	AA133312	Arabidopsis thalia
41	117	4.2	565	21	AA146726	Arabidopsis thalia
42	116	4.2	427	21	AA146725	Arabidopsis thalia
43	116	4.2	502	21	AA146724	Arabidopsis thalia
44	116	4.2	552	21	AA146723	Arabidopsis thalia
45	113	4.1	407	19	AA11034	H. pylori ORF 04ep

ALIGNMENTS

RESULT 1
ID AAY40901 standard; Protein: 538 AA.
XX AC AAY40901;
XX AC AAY40901;
DT 18-JAN-2000 (first entry)
XX Arabidopsis thaliana Na/H transporter ATNHX1.
XX Sodium; proton; antipor; transporter; salt tolerance; salt management;
KW transgenic plant; survival; soil; farming; accumulation; irrigation;
KW crop.
XX Arabidopsis thaliana.
XX Arabidopsis thaliana.
PN WO9947679-A2.
XX WO9947679-A2.
PD 23-SEP-1999.
XX 23-SEP-1999.
XX 18-MAR-1999; 99WO-CA00219.
XX 18-MAR-1999; 98US-0078474.
PR 15-JAN-1999; 99US-0116111.
XX (BLUM/) BLUMWALD E.
PA (APSE/) APSE M.
PA (SNED/) SNEDDEN W.
PA (AHAR/) AHARON G.
XX Blumwald E, Apse M, Snedden W, Aharon G;
XX WPI; 1999-571840/48.
DR N-PSDB; AA222591.

XX PT Nucleic acid molecules encoding sodium/proton transport polypeptides,
 XX useful in genetic engineering salt tolerance in crop plants -
 PS Claim 36; Fig 1A; 93pp; English.
 XX The invention relates to an isolated nucleic acid molecule encoding
 CC a plant Na/H antiporter (PNHX) transporter polypeptide, or a fragment
 CC and capable of increasing salt tolerance in a cell. This sequence
 CC corresponds to the AtNHX1 transporter from Arabidopsis thaliana.
 CC The Na/H transporter polypeptides provide a means of intracellular
 CC salt management, particularly in plants. The sequences are useful for
 CC producing transgenic plants that are capable of surviving in soil with
 CC high salt levels that would normally inhibit growth of the crop species.
 CC This would be useful in farming land in areas that are generally
 CC considered unproductive through salt accumulation and poor irrigation,
 CC e.g. in India, Australia, and prairies in USA or Canada. Commercial
 CC crops, such as potatoes, tomatoes, brassica, cotton, sunflower,
 CC strawberries, spinach, lettuce, rice, soybean, corn, wheat, rye, barley,
 CC atriplex, sorghum, alfalfa, salicornia and others would benefit from
 CC increased salt tolerance.
 XX SQ Sequence 538 AA;

Query Match 100.0%; Score 2755; DB 20; Length 538;
 Best Local Similarity 100.0%; Pred. No. 3.9e-282;
 Matches 538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MLDLSVSKLPSTSDHASVVALNFVALLCACIVLGHLEENRMNSESITALLIGLGTG 60
 Db 1 mldlsvsklpstsdhasvvalnfvallcacivlgllleenrwmnesitalligltg 60
 QY 61 VTILLISKGKSHLLVTFSEDLFFIYLLPPIIFNAGFQVKKKOFFRNFVTIMLFGAVGTII 120
 Db 61 vtilliskgkshllvtfseddlffiyllppiiifnagfqvkkkoffrnfvtimlfgavgtii 120
 QY 121 SCTIISLGVTQFFKKLDIGTFDGLDYLAIGAIFAAPDSVCTLOVFNQDETPLLVSFVGE 180
 Db 121 sctiislgvtqffkkldigtfdglgdylaigaifaatdsdvctllqvnqdetppllvsfvg 180
 QY 181 GYVNDATSVVFNFAIQSFDTLHNEAAPHLLGNFLYFLSTLLCAATGLISAVYIKKL 240
 Db 181 gyvndatsvvfnfaiaqsfdtlhneaaphllgnfllyflstllgaatglisavyyikkl 240
 QY 241 YFGRHSTDREVALMMLMAYLSYMLAEFLDLSGLTAVFFCGIWMSHYTHWNTVTESSRITTK 300
 Db 241 yfgrhstdrevalmmlmaylsymlaelfldlsgltavffcgivmshytnhntvtesrirttk 300
 QY 301 HTFATLSPLAETFTFLYVGMADLIDKWRVSPTPCTSIATSVSSILMGLVMVGRAAFVPL 360
 Db 301 htfatlsplaetftflvygmaldidkwrvsptpctsiatssilmglvmvgraafvpl 360
 QY 361 SFLSNLAKKNGSEKINFNMQVVIWWSGLMRGAVSMALAYNKFTFRAGHTDVRGNALMITST 420
 Db 361 sflsnlakkngsekinfnmqvviwsglmrgavsmalaynknkftfraghtdvrgnalmitst 420
 QY 421 ITVCLFSFVPGMLTKPLISYLLPHONATWSMLSDNTPKSTHPLDQDSFIEPSGNHN 480
 Db 421 itvclfsfvpgmltkplisyllphonatwsmlsdntpksthplldqdsfiepsgnhn 480
 QY 481 VPRPDSIRGFLTRPTRTVHYVWQFDDSFMRPVFGGREGVFPVPGSPTRNPPDLASKA 538
 Db 481 vprpdsirgfltrptrtvhyvwqfddsfmrpvfggregvfpvpgsptrnppdlaska 538

RESULT 2

AA40905
 ID AA404905 standard; Protein: 547 AA.
 XX
 AC AA40905;
 XX

DT -18-JAN-2000 (first entry)

XX DE Arabidopsis thaliana Na/H transporter.
 XX Sodium; proton; antiporter; transporter; salt tolerance; salt management;
 KW transgenic plant; survival; soil; farming; accumulation; irrigation;
 KW crop.
 XX Arabidopsis thaliana.
 XX WO9947679-A2.
 PN 23-SEP-1999.
 XX 18-MAR-1999; 99WO-CA00219.
 XX 18-MAR-1998; 98US-0078474.
 PR 15-JAN-1999; 99US-0116111.
 XX {BLUM/} BLUMWALD E.
 PA {APSE/} APSE M.
 PA {SNED/} SNEDDEN W.
 PA {AHAR/} AHARON G.
 XX BLUMWALD E, APSE M, Snedden W, Aharon G;
 PI WPI; 1999-571840/48.
 DR N-PSDB; AA222595.
 XX Nucleic acid molecules encoding sodium/proton transport polypeptides,
 PT useful in genetic engineering salt tolerance in crop plants -
 PS Claim 50; Fig 5A-B; 93pp; English.
 XX The invention relates to an isolated nucleic acid molecule encoding
 CC a plant Na/H antiporter (PNHX) transporter polypeptide, or a fragment
 CC and capable of increasing salt tolerance in a cell. This sequence
 CC corresponds to a transporter from Arabidopsis thaliana.
 CC The Na/H transporter polypeptides provide a means of intracellular
 CC salt management, particularly in plants. The sequences are useful for
 CC producing transgenic plants that are capable of surviving in soil with
 CC high salt levels that would normally inhibit growth of the crop species.
 CC This would be useful in farming land in areas that are generally
 CC considered unproductive through salt accumulation and poor irrigation,
 CC e.g. in India, Australia, and prairies in USA or Canada. Commercial
 CC crops, such as potatoes, tomatoes, brassica, cotton, sunflower,
 CC strawberries, spinach, lettuce, rice, soybean, corn, wheat, rye, barley,
 CC atriplex, sorghum, alfalfa, salicornia and others would benefit from
 CC increased salt tolerance.
 XX SQ Sequence 547 AA;

Query Match 91.3%; Score 2516; DB 20; Length 547;
 Best Local Similarity 100.0%; Pred. No. 7e-257;
 Matches 496; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MLDLSVSKLPSTSDHASVVALNFVALLCACIVLGHLEENRMNSESITALLIGLGTG 60
 Db 1 mldlsvsklpstsdhasvvalnfvallcacivlgllleenrwmnesitalligltg 60
 QY 61 VTILLISKGKSHLLVTFSEDLFFIYLLPPIIFNAGFQVKKKOFFRNFVTIMLFGAVGTII 120
 Db 61 vtilliskgkshllvtfseddlffiyllppiiifnagfqvkkkoffrnfvtimlfgavgtii 120
 QY 121 SCTIISLGVTQFFKKLDIGTFDGLDYLAIGAIFAATDSVCTLOVFNQDETPLLVSFVGE 180
 Db 121 sctiislgvtqffkkldigtfdglgdylaigaifaatdsdvctllqvnqdetppllvsfvg 180
 QY 181 GYVNDATSVVFNFAIQSFDTLHNEAAPHLLGNFLYFLSTLLCAATGLISAVYIKKL 240
 Db 181 gyvndatsvvfnfaiaqsfdtlhneaaphllgnfllyflstllgaatglisavyyikkl 240
 QY 241 YFGRHSTDREVALMMLMAYLSYMLAEFLDLSGLTAVFFCGIWMSHYTHWNTVTESSRITTK 300

```

|||||
Db 241 yfgrstdevalmmlmaylsymlaelfalsgiltvfcgvmshytnvntesrittk 300
QY 301 HTFATLSFLAETFIIFYVGMDALDIDKWRVSVDTPGTSIAVSSILMGLVMVGRAPFVPL 360
Db 301 htfatlsflaetfifyvgmdaldidkwrsvsdtpgtsiavsilmlglvmvgrafvpl 360
QY 361 SFLSNLAKKNQSEKINFNMVQVWISGLMRGAVSNALAYNKFTRAGHTDVRGNAMITST 420
Db 361 sflsnlakknqsekinfnmqvwwsglmrgavsmalaynkftraghtdvrgnalmist 420
QY 421 ITVCLFSTVFGMLTKPLISYLLPHQNATTSMLSDNTPKSIHIPLLDQDSFIIEPISGNHN 480
Db 421 itvclfstvfgmltkplisyllphqnattsmldndtpksihppllqdsfiiesgnhn 480
QY 481 VPRPDSIRGFLRPRTR 496
Db 481 vprpdsirgflrprtr 496

RESULT 3
AAB12786
ID AAB12786 standard; Protein; 555 AA.
XX
AC AAB12786;
XX
DT 23-NOV-2000 (first entry)
XX
DE AtIiplex gmelini Na+ and H+ antiporter protein.
XX
KW AtIiplex gmelini; Na plus and H plus antiporter protein;
KM Na+ and H+ antiporter protein; transformed plant; high salt tolerance.
XX
OS AtIiplex gmelini.
XX
PN JP2000157287-A.
XX
PD 13-JUN-2000.
XX
PF 16-SEP-1999; 99JP-0261606.
XX
PR 24-SEP-1998; 98JP-0269504.
XX
PA (SHOK-) SHOKUBUTSU KOGAKU KK.
XX
DR WPI; 2000-468209/41.
DR N-PSDB; AAA72926.
XX
PT An Na+ and H+ plus antiporter protein and a gene encoding it
XX
PS Disclosure; Page 10-12; 16pp; Japanese.
XX
CC The present sequence represents an Na+ and H+ antiporter protein
CC isolated from AtIiplex gmelini. The Na+ and H+ antiporter protein and
CC gene encoding it are useful for the preparation of transformed plants
CC with high salt tolerance, e.g. for growth in arid land.
XX
SQ Sequence 555 AA;

Query Match 78.7%; Score 2168.5; DB 21; Length 555;
Best Local Similarity 77.7%; Pred. No. 3.7e-220;
Matches 425; Conservative 53; Mismatches 56; Indels 13; Gaps 5;

QY 2 LDSLVS-KLPSTSDSHASVWALNLFVALLCIVGLHLEENRWMSITALLIGLGTG 60
Db 5 lssllsgkmdalttsdhasvsmmlfvallcgvighleerwnmesitalliglatg 64
QY 61 VTILLISGKSHLLVFSEDLFFIYLLPPIIFNAGFQVKKQFFRNFTIMLFAGAVGTII 120
Db 65 vvilllsgkshllvfseddlffiyllppifnagfqvkkkqffrnftivlfgavgtlv 124
QY 121 SCTIISLGVTQFPKKLDIGTFDGLDYLAIGAIFATDTSVCTQLQVNLQDETPLLVSYPGE 180

```

```

|||||
Db 125 sftlslgalsifkklidgclleadylaigaifaatdsvctqivngdetpilsylvfe 184
QY 181 GVVNDATSVVFNATQSFDTLHNHEAAFLHNLGNFLYLFLSTLIGATGLISAYVKKL 240
Db 185 gvvndatsvvlfnatqsfldtridhrialqfmgfnflyfiastilgaftgilsaylikkl 244
QY 241 YGRHSTDEVALMMLMAYLSYMLAELPDLSGILTVPFCGIVMSHYTHWNTESRITTK 300
Db 245 yfgrstdevalmmlmaylsymlaelfalsgiltvfcgvmshytnvntesrvttk 304
QY 301 HTFATLSFLAETFIIFYVGMDALDIDKWRVSVDTPGTSIAVSSILMGLVMVGRAPFVPL 360
Db 305 hafatlsfvaefvflyvgmdaldiekwrfsdpsgiasvavssillglvmvgrafvpl 364
QY 361 SFLSNLAKKNQSEKINFNMVQVWISGLMRGAVSNALAYNKFTRAGHTDVRGNAMITST 420
Db 365 swlmnfakksqekvtngqglvwwaglmrgavsmalaynqftrsghtqlrignalmist 424
QY 421 ITVCLFSTVFGMLTKPLISYLLPHQNATTSMLSDNTPKSIHIPLLD--QDSFIE-P 475
Db 425 isvlfstmvfglltkplimfilpqpkhftscstsvdvgspksyslpllegnqgdevdvg 484
QY 476 SGNH-----NVPRPDSIRGFLRPRTRFTVHYWRQFDDSPMRPVFGRGVPVPGSPT 528
Db 485 ngnhedttrptivirpssirmalnaphthvhwkfkddsmrpvfgrgvfpvpgspt 544
QY 529 ERNPPDL 535
Db 545 eqstnl 551

RESULT 4
AAB73253
ID AAB73253 standard; Protein; 553 AA.
XX
AC AAB73253;
XX
DT 14-MAY-2001 (first entry)
XX
DE Protein regulating the pH of vacuoles.
XX
KW Vacuole pH regulation; flower colour.
XX
OS Nierembergia hybrida.
XX
PN WO200114560-A1.
XX
PD 01-MAR-2001.
XX
PF 24-AUG-2000; 2000WO-JP05722.
XX
PR 24-AUG-1999; 99JP-0236800.
XX
PA (SUNR ) SUNTORY LTD.
XX
PI Iida S, Tanaka S, Inagaki Y;
XX
DR WPI; 2001-191648/19.
DR N-PSDB; AAF75765.
XX
PT Morning glory-originated gene encoding a protein with pH regulation
PT activity in vacuoles, useful in controlling flower color to give new
PT breeds of colorful plants for cut flowers, particularly applicable in
PT horticulture -
XX
PS Example 7; Page 50-52; 68pp; Japanese.
XX
CC The present sequence is a protein, which has vacuolar pH regulatory
CC activities. The protein enables flower colour to be controlled via
CC regulation of the vacuolar pH; colours can range from blue to red in
CC colour spectrum. The protein is useful in controlling flower colour to
CC give new breeds of colourful plants for cut flowers, particularly

```

cc applicable in horticulture.

XX
50 Sequence 553 AA;

Query Match	75.9%;	Score 2091.5;	DB 22;	Length 553;
Best Local Similarity	76.3%;	Pred. No. 4.9e-212;		
Matches 405:	Conservative	53;	Mismatches 72;	Indels 1;

Qy	4	SLVSKLPSTSOHSAVVALNLFVALLCACIVLGHLEENRWNESITALLIGLGTGVTI	63
Db	7	tlgkmmnlittdhqsqvsvnlfvallcacivlghleerwnmesitalvsgstgwl	66
Qy	64	LLISGKSSHLVFSDELPIYLPLIFNAGQVKKQFFRNFVTIMLFAGVAGTIIIST	123
Db	67	llisggkshlrvfsedlffiyllppifnaqfvkksffrnfstimifgavgtlslfi	126
Qy	124	IISLGVTOFPKKLDIGPFDLGDYLATGATPAANDSVCTLOVLNQDFTPLLYSLVFGEGV	183
Db	127	lissagaifkkmldighleigdylaigaifaatdsvtclqvlnqeetpillyslvfgvgv	186
Qy	184	NDATSVVFNAAIQSFDTLNLHHEAAPHLLGNFYLYLLSTLLCAAAGLSAVVKKLYFG	243
Db	187	ndatsvfvnagvqfshstskgalkignfiylfasstfigavglisafiikklyfg	246
Qy	244	RHSTDREVALMMLAYLSYMLAELFDLSGILTVFFCGIVMSHTWHNVTESSRITTKHTF	303
Db	247	rhtsdrevalmmlaylsymlaelfylsgiltvfcgivmshywhnvteessrvtkhtf	306
Qy	304	ATLSFLAETFTFLVYGMDALIDKWRVSPTPGTSTAVSSILMGLVMVGRAAFVPLSTL	363
Db	307	atlsflaeftflvygmdaldiekwkvfsdsgptcsikvssillglvlvrgatvflpsl	366
Qy	364	SNLAKKNQSEKINFNMQVYIWWSGLMRGAVSMALAYNKFTFRAGTDDVRGNAMITGTV	423
Db	367	snlktknpedkisinfaqvqtiwagvlmrgavsmalaynqftrggtqlranaimitstiv	426
Qy	424	CLFSTIVFGMLTKPLISVLLPHQNAATYSMLSDD-NTPKSTHILPLLODSDSIEPSSGNHNPV	482
Db	427	vlfstvwfgmltkplllllllsqkhlrmisemtpmkfsfvlpltdqdsadlgrhvp	486
Qy	483	RPDSLRGFLTPTRTVHYWRFQDSDSMRPVFGRGFVFPVPGSPTEPNPP	533
Db	487	rbsjrlmllsrnebtvhwvwrkfdnarmvryfgarqfvafvoosptepvcp	537

RESULT 5
AAB73252
ID AAB73252 standard: protein: 540 AA.

AAC AAB73252:

XX
XX

DT 14-MAY-2001 (first entry)

XX

DE Protein regulating the pH of vacuoles

XX
 12-10-21 11:00 AM

УУ
КУ
МУ
и
вас
и
те

OS Petunia hybr.

XX

PN WO200114560-A1.

[illegible]

FD 01-MAR-2001.

24-AUG-2000 2000W0-

N-PSDB: AAF75764.

XX Morning glory-originated gene encoding a protein with pH regulation
PT activity in vacuoles, useful in controlling flower color to give new
PT breeds of colorful plants for cut flowers, particularly applicable in
PT horticulture -

Example 7: Page 42-45; 68pp; Japanese.

The present sequence is a protein, which has vacuolar pH regulatory activities. The protein enables flower colour to be controlled via regulation of the vacuolar pH, colours can range from blue to red in colour spectrum. The protein is useful in controlling flower colour to give new breeds of colourful plants for cut flowers, particularly applicable in horticulture.

AA	Sequence	540 AA;
50		

Query Match	74.8%;	Score 2059.5;	DB 22;	Length 540;
Best Local Similarity	75.3%;	pred. No. 1.1e-208;		
Matches 397:	Conservative	57;	Mismatches 72;	Indels 1;

Qy	4	SLVSKLPSLSTSDHSAVVALNLFVALLCACIVLGHLEENRWNNESITALLILGLGTVTI	63
Db	7	tl gnvdr stsdhqs vsnlf vallicaciv ghleer wnmesit alv igsc t v l	66
Qy	64	LLISKGSSHLILVFESDEFFIYLLPPIIFNAGQVKKQKOFFRNFVTIMLFAGVGTIISCT	123
Db	67	llisggnshlilvfesd ffi y llp i f n a g q v k k o f f r n f v t i m l f a g v g t i s c t	126
Qy	124	IISIGVQTFPKKLDIGTFDGLDYLAICAIFAAATDSVCTLOVNODETPLLYSLVFGEVGV	183
Db	127	i s i g i g a f i k m i n g i s l e g d y i a i g i a f a a t d s v c t l o v n o d e t p l l y s l v f g e v g v	186
Qy	184	NDATSVVVFNAIQSFDLTHLNEHAFFLLGNFLYFLFLLSTLLGAATGLISAYIKKLYFC	243
Db	187	ndat sv v f n a i q s f d l h l n e h a f f l l g n f l y f l l s t l l g a a t g l i s a y i k k l y f c	246
Qy	244	RHSTDREVALMWLMAYLSYMYLAELFDLSGILTVFFCGIVMSHYTHWNVTSSRLTYKHTF	303
Db	247	r h s t d r e v a i l m a y l s y m y l a e l f d l s g i l t v f f c g i v m s h y t h w n v t s s r l t y k h t f	306
Qy	304	ATLSPLAETFFLVXGMDALDIDKWRSVSDPPGNSIAVSSILMGLVMVGRAAFVPLLSFL	363
Db	307	a t l s p l a e t f f l v x g m d a l d i d k w r s v s d p p g n s i a v s i l m g l m v m g r a a f v p l l s f l	366
Qy	364	SNLAKKNSEKINFNMQVVIWWSGLMRGAVSMALAYKNKTRAGHTDVRGNAMITSTIV	423
Db	367	s n l a k k n s e k i n f n m q v v i w w s g l m r g a v s m a l a y k n k t r a g h t d v r g n a m i t s t i v	426
Qy	424	CLFSTVVFGLMKPLLSLLLP-HQNAATFSMLSDNTPTKS IHPLLDQDSFTEPSSGNHVP	483
Db	427	v l f s t v v f g l m c k p l l s l l p - h q n a t f s m l s d n t p t k s i h p l l d q d s f t e p s g h v p	486
Qy	483	RPDSIRGETRTPTRVHYWQFDSDSEMRPFVGGGFVFPVGSPT 529	
Db	487	r p d s i r g e t r p t r v h y w q f d s d e m r p f v g g g f v p p g s p t 533	

RESULT 6
AAB73251
ID AAB73251 standard: protein: 542 AA:

AA
AC
AAR73251:

2000

DT 14-MAY-2001 (first entry)

XX

DE Protein regu

XX WO200114560-A1.
PN 01-MAR-2001.
XX 24-AUG-2000; 2000WO-JP05722.
XX 24-AUG-1999; 99JP-0236800.
XX (SUNR) SUNTORY LTD.
XX Tida S, Tanaka S, Inagaki Y;
XX WPI; 2001-191648/19.
DR N-PSDB; AAF75752.
XX Morning glory-originated gene encoding a protein with pH regulation
PT activity in vacuoles, useful in controlling flower color to give new
PT breeds of colorful plants for cut flowers, particularly applicable in
PT horticulture -
XX
PS Claim 2; Page 30-33; 68pp; Japanese.
XX
XX The present sequence is a protein from Morning Glory, which has vacuolar
CC pH regulatory activities. The protein enables flower colour to be
CC controlled via regulation of the vacuolar pH, colours can range from blue
CC to red in colour spectrum. The protein is useful in controlling flower
CC colour to give new breeds of colourful plants for cut flowers,
CC particularly applicable in horticulture.
XX
SQ Sequence 542 AA;

Query Match 74.1%; Score 2041; DB 22; Length 542;
Best Local Similarity 73.8%; Pred. No. 1e-206;
Matches 399; Conservative 57; Mismatches 71; Indels 14; Gaps 4;
QY 2 LDSLVSKLPSSL-STSASHVVALNLFVALLCACIVLGHLEENRWNNESITALLIGLGTG 61
DB 5 lssallqc-sdlftsdhasvsmnlfvallicacivghlleenrwnesitalligltcv 63
QY 62 TILLISKGSSHLVSEDLFFIYLLPPIIFNAGFOVKKQFFRNFVTIMLFGAVGTII 121
DB 64 vlllsgksshlvseDLFFIYLLPPIIFNAGFOVKKQFFRNFVTIMLFGAVGTII 123
QY 122 CTIISLGVTQFFKKLDIGTFDGLDYLAIGAIFAATDSVCTLOVNOQDETPLLYSLVFGE 181
DB 124 csaisfgavkifkhlidldfgydlaigaifaatdsVCTLOVNOQDETPLLYSLVFGE 183
QY 182 VVNDATSVVVFNAIOFSLTHLNHEAAPHLLGNFLYLFLLSTLLGAATGLISAYVIRKLY 241
DB 184 vvndatsvvlfnagisdmtsdtpkiglfnflyflsftlglvlgicaylikly 243
QY 242 FGRHSTDREVALMMLMAYLSYMLAELFDLSGLTVPFFCGIVMSHYTHWNTVESSRITTK 301
DB 244 fgrhstdrevalmmlmsylsymlaelflylsgltvffcgivmshytwntveSSRITTK 303
QY 302 TFATLSLAETFIYLVGMDALDIDKWSVSDTPCTSIASVSSILMGLVMVGRAAFVPLS 361
DB 304 sfatlsfvaetfiylvgmdalidiekvfvknsqglsvavssilvgllvgraaVPLS 363
QY 362 FLSNLAKNQSKINFNMQVITWNSGLMRGAVSMALAYNKETFRAGHTDVRGNAMITSTI 421
DB 364 flsnlakknsdkisfrqiiilwaglmrgavslalaynkfttsghtshenaimstv 423
QY 422 TVCLSTVVEGMLTKPLISYLLP-----HQNATTSMLEDDNTPKSIHPLDQ--DS 471
DB 424 tvclstvviglmTKPLINILPpKmpsgshsmnt---sepspkhtvplldngpds 480
QY 472 FIEPSGNHNVPRPDSIRGFLTRPTRTVHYWRFQDPSMRPVFEGRGVFPVPGSPERN 531
DB 481 esdmigpeavptalmllrtpthtvhywrkfkfadsfmrpvfgrgfvfvgpsvqgs 540

QY 532 P 532
DB 541 p 541
RESULT 7
AAB11832
ID AAB11832 standard; Protein; 535 AA.
XX
AC AAB11832;
XX
DT 07-NOV-2000 (first entry)
XX
DE Rice Na+/H+ antiporter, OsnHX1.
XX
KW OsnHX1; Na+/H+ antiporter; sodium/proton antiporter; countertransporter;
PT active transport; rice; transgenic plant; salt-tolerance.
XX
OS Oryza sativa.
XX
PN WO200037644-A1.
XX
PD 29-JUN-2000.
XX
PF 22-DEC-1999; 99WO-JP07224.
XX
PR 22-DEC-1998; 98JP-0365604.
XX
PA (NORQ) JAPAN MIN AGRIC FORESTRY & FISHERIES.
XX
PI Fukuda A, Tanaka Y;
XX
DR WPI; 2000-442672/38.
DR N-PSDB; AAA61876.
XX
PT DNA encoding a sodium ion and proton counter-transporter protein of
PT rice origin for production of salt tolerant rice transformants -
XX
PS Claim 1a; Fig 1; 43pp; Japanese.
XX
CC This sequence represents a novel rice Na+/H+ antiporter
CC (countertransporter), OsnHX1. The invention relates to OsnHX1 and
CC nucleic acids which encode it; vectors, host cells and transgenic plants
CC containing OsnHX1 nucleic acids; recombinant expression of OsnHX1; and
CC antibodies which recognise OsnHX1. OsnHX1 nucleic acids are useful in the
CC production of salt tolerant transgenic plants.
XX
SQ Sequence 535 AA;

Query Match 73.4%; Score 2023.5; DB 21; Length 535;
Best Local Similarity 73.3%; Pred. No. 7.2e-205;
Matches 389; Conservative 60; Mismatches 79; Indels 3; Gaps 2;
QY 2 LDSLVSKLPSSL-STSASHVVALNLFVALLCACIVLGHLEENRWNNESITALLIGLGTG 60
DB 3 mevaaralgalttsdyasvsnlfvallicacivghlleenrwnesitalligltcv 62
QY 61 VTILLISKGSSHLVSEDLFFIYLLPPIIFNAGFOVKKQFFRNFVTIMLFGAVGTII 120
DB 63 vvilmtckgsshlvseDLFFIYLLPPIIFNAGFOVKKQFFRNFVTIMLFGAVGTII 122
QY 121 SCTIISLGVTQFFKKLDIGTFDGLDYLAIGAIFAATDSVCTLOVNOQDETPLLYSLVFGE 180
DB 123 sftfistaaiafsmmigtldvgdflaigaifatsdsvctlgvlnqdetpflysvfge 182
QY 181 GVNDATSVVVFNAIOFSLTHLNHEAAPHLLGNFLYLFLLSTLLGAATGLISAYVIRKLY 240
DB 183 gvndatsvvlfnaglnqndvlhdaavvklfnfflyflsftlglvfgagilkl 242
QY 241 YFGRHSTDREVALMMLMAYLSYMLAELFDLSGLTVPFFCGIVMSHYTHWNTVESSRITTK 300
DB 243 ygrhstdrevalmmlmaylsymlaelldlgltvffcgivmshytwntveSSRITTK 302

Qy	301	HTFATISFLAETFIYVGMALDIDKWRVSVDTPGTSIAVSSILMLGVWGRAAFVPL	360
Db	303	hafatisflaetfiyvgmdaldiekewafadrgpgksigissillllvlgiraafvpl	362
Qy	361	SFLSNLAKKNQSEKINFNQVVIWWSGLMRGAVSALAYNKETRAGHDVGRNALMITST	420
Db	363	sfislnltkapnekikwrgdgvllwaglmrgavsalaynkfrsghtqqlhgnalmist	422
Qy	421	ITVCLFTVVFGMLTPLYLLPHQNATTSMLSDDNTPKSIHPLDQDSFIPIPSGNHN	480
Db	423	itvvlfstvmfgmttkplrlilpasghpyt--sepsapkslhspllismagdslestin	480
Qy	481	VPRPDSIRGFLTRPTVTHYYWRQEDDSFMRPVFGGRGVFVPGSPSTERN	531
Db	481	ivresslmltlktotbvwwrkfdalrmomfaqrqfvfsspsleqas	531

RESULT	8	
AAB73254		
ID	AAB73254	standard; Protein; 555 AA.
XX		
XX	AC	
XX	AAB73254;	
XX		
DT	14-MAY-2001	(first entry)
XX		
XX		
DE		Protein regulating the pH of vacuoles.
XX		
XX		
KW		Vacuole pH regulation; flower colour.
XX		
XX		
XX		Torenia hybrida.
XX		
PN	WO200114560-A1.	
XX		
PD	01-MAR-2001.	
XX		
PF	24-AUG-2000;	2000WO-JP05722.
XX		
PR	24-AUG-1999;	99JP-0236800.
XX		
XX		
PA	(SUNR)	SUNTORY LTD.

	Query Match	72.0%;	Score 1982.5;	DB 22;	Length 555;
	Best Local Similarity	73.1%;	Pred. No. 1.6e-200;		
	Matches 391;	Conservative 53;	Mismatches 80;	Indels 11;	Gaps 3;
QY	7	SKLP	SLSTSDSHASVVALNLFVALLCACIVLGHLLLENRMWNESTALLIGTGTVILLI	66	
	l:	: :	l : l l l l l	l l l l l	l l l l l
Db	12	setdnlwssghgsvvailtvtllctcivighlleenrwmnesilaligatgvillli	71		
OY	67	SKGKSSHLLVFSEDLFFYVLLPPIIFNAGFOVKKKOFFERNFVTIMFGAVGTITISCTTIS	126		

[illegible]

OS Saccharomyces cerevisiae.
 PN WO9947679-A2.
 XX 23-SEP-1999.
 XX 18-MAR-1999; 99WO-CA00219.
 XX 18-MAR-1998; 98US-0078474.
 PR 15-JAN-1999; 99US-0116111.
 XX (BLUM/) BLOWALD E.
 PA (APSE/) APSE M.
 PA (SNED/) SNEDDEN W.
 PA (AHAR/) AHARON G.
 XX Blumwald E, Apse M, Snedden W, Aharon G;
 DR WPI; 1999-571840/48.
 XX Nucleic acid molecules encoding sodium/proton transport polypeptides,
 PT useful in genetic engineering salt tolerance in crop plants -
 XX Disclosure; Fig 8B; 93pp; English.
 XX The invention relates to an isolated nucleic acid molecule encoding
 CC a plant Na/H antiport (PNHX) transporter polypeptide, or a fragment
 CC and capable of increasing salt tolerance in a cell. This sequence
 CC corresponds to a Na/H antiport transporter from the budding yeast
 CC Saccharomyces cerevisiae. The Na/H transporter polypeptides provide
 CC a means of intracellular salt management, particularly in plants. The
 CC sequences are useful for producing transgenic plants that are capable
 CC of surviving in soil with high salt levels that would normally inhibit
 CC growth of the crop species. This would be useful in farming land in areas
 CC that are generally considered unproductive through salt accumulation and
 CC poor irrigation, e.g. in India, Australia, and prairies in USA or Canada.
 CC Commercial crops, such as potatoes, tomatoes, brassica, cotton,
 CC sunflower, strawberries, spinach, lettuce, rice, soybean, corn, wheat,
 CC rye, barley, triplic, sorghum, alfalfa, salicornia and others would
 CC benefit from increased salt tolerance.
 XX Sequence 533 AA;
 SQ

Query Match 20.5%; Score 564; DB 20; Length 633;
 Best Local Similarity 26.8%; Pred. No. 1.4e-50;
 Matches 164; Conservative 121; Mismatches 200; Indels 128; Gaps 21;

QY 14 TSDHASVVALNLFVALLCACIVLGHLEEN--RWNNESITALLIGLGTGVILLISKGS 71
 DB 55 teemfswalfmllilalswssyyltqkriravhetvisifgmviglirmspghyi 114
 QY 72 SHLVFSEDLFFIYLLPPIIFNAGFOVKKQFRNFVITMLFGAVGTIISCTIISLGVQ 131
 DB 115 qdvtfnssyffnvlppilnsgyelncqvnfnmllsillfaipgttisavvigi-ily 173
 QY 132 FPKKLDIGFDL--GDYLAIGAIFAATDSVCTLOVLNODER-PLLYSLVFGGVVNDATS 188
 DB 174 iwtflgesidifadamsvgatlsatpvtllsifnkyvdpklytlifgeslndais 233
 QY 189 VVVFNAIQSPDLPHLNEHAAPHLGNFLFLSTLLGAATGLISAYVIRKLYFGRHSTD 248
 DB 234 ivmfetcqkthgqpatfssvfegagiflmtfsvlligvlgilvallkthirry-pq 292
 QY 249 REVALMMLAYLSYMLAEFLDLISGLITVFECGIVMSHYTHWNTVSSRITTKHFAFLSP 308
 DB 293 iesclillayesffnsngchmsgvislfcglttkhyyanmsrrsrtikylfqlar 352
 QY 309 LAETFFLYVGMDALDIDKWSRSDTPGTS-----IAVSSILMGLVMVGRAAFVFP 359
 DB 353 lsenfifiylglef-----levelvykplliivaal---sicvawcavfp 396
 QY 360 LS-----FLSNLAKNQS--EKINFNQVVIWWSGLMRGAVSWALAYN---- 400

DB 397 lsqfwnwiyrvktirsmgitsgenisvdpelpynqmntfwagl-rgevvalalgiqe 455
 QY 401 -KFRAGHTDVRGNAMITSTITVCLFSTVVGMLTKPLISYLLPHQNTATMSLSDDT- 458
 DB 456 ykft-----llatlvvvvltviifggttagmlevl----niktgciseads 499
 QY 459 -----PKSTHI-----PLLDQDSFIEPSGNHNVPRPDSIRGFLT 492
 DB 500 ddefdieapraanllngssiqtdlpgysdnnspsididqf-avssnknlpnnisttgn 558
 QY 493 -----RPRTV-----HYWROFDSFMRPVGGRGFPFV 523
 DB 559 fgglnetentspnparsmdknrlrklgtifnsdsgwfgnfdqevlqpvfld-nvpsl 617
 QY 524 PGSPTERNPDL 536
 DB 618 qdsatq-spadsf 629

RESULT 13
 AAB90555
 ID AAB90555 standard; Protein; 509 AA.
 XX
 AC AAB90555;
 XX
 DT 01-JUN-2001 (first entry)
 XX
 DE Human secreted protein, SEQ ID NO: 93.
 XX
 KW Human; secreted protein; immunomodulatory; antisclerotic;
 KW dermatological; antiinflammatory; anti-HIV; cytostatic; cardiac;
 KW vascular; anti-angiogenic; ophthalmological; neuroprotectant;
 KW neoplastic; anticonvulsant; antiaizhelmers; antiparkinsonian;
 KW antimicrobial; vulnary; vaccine; gene therapy; cancer;
 KW protein coordinate data; infection.
 XX
 OS Homo sapiens.
 XX
 PN WO200121658-A1.
 XX
 PD 29-MAR-2001.
 XX
 PF 22-SEP-2000; 2000WO-US26013.
 XX
 PR 24-SEP-1999; 99US-0155709.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ni J, Baker KP, Birse CE, Ebner R, Fiscella M, Komatsoulis GA;
 PI Lafleur DW, Moore PA, Olsen HS, Rosen CA, Ruben SA, Soppet DR;
 PI Young PE, Wei P, Florence KA;
 XX
 DR WPI; 2001-235311/24.
 DR N-PSDB; AAF97895.
 XX
 PT Nucleic acids encoding 32 human secreted polypeptides, useful for
 PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's
 PT disease and diabetic retinopathy -
 XX
 PS Claim 11; Page 779-780; 890pp; English.
 XX
 CC The present sequence is one of 32 novel human secreted polypeptides. The
 CC nucleic acid molecules and polypeptides may be used in the prevention,
 CC diagnosis and treatment of diseases such as immune disorders (e.g.
 CC multiple sclerosis, systemic lupus erythematosus and human
 CC immuno-deficiency virus (HIV) infections), hyperproliferative disorders
 CC (e.g. cancers and Gaucher's disease), cardiovascular diseases
 CC (e.g. scimitar syndrome, Chaga's cardiomyopathy and coronary
 CC arteriosclerosis), angogenic disorders (e.g. corneal graft
 CC neovascularisation and diabetic retinopathy), neurological disorders
 CC (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease),
 CC infectious diseases and/or for promoting wound healing, regeneration

CC and/or chemotaxis. The nucleic acid molecules may be used to produce the
 CC secreted polypeptides. They may also be used as DNA probes in diagnostic
 CC assays to detect and quantitate the presence of similar nucleic acid
 CC sequences in samples. The polypeptides may be used as antigens in the
 CC production of antibodies and in assays to identify modulators of
 CC their expression and activity.

XX Sequence 509 AA;

Query Match 20.4%; Score 563; DB 22; Length 509;
 Best Local Similarity 31.1%; Pred. No. 1.3e-50;
 Matches 162; Conservative 90; Mismatches 165; Indels 104; Gaps 18;

QY 21 VALNLFVALLCIVLGHLEEN--RMNSESITALLIGLGTGV----- 61
 DB 19 vxlltfillltltltlwtlkrvrrfthetglamyglivgvlrygtgtpatsgrdksls 78
 QY 62 -----TILL-----ISKSSHL-----LVFSEDLFFIYLLPPI 90
 DB 79 ctqdrastltnvnsqkffeytlkgeispgkinsveqndmrkvcfdpevfnnallppi 138
 QY 91 IFNAGVQKKQFFRNFVTIMLFGAVGTIISCTIIS---LGVTFQFFKLDIGTFDLGD-- 145
 DB 139 ifhagsyikkrhfrnlgslayafgtaxscflignlmvgvklmklmg----qlsdf 194
 QY 146 ---YLAIGALFAATDSVCTQLVNLQDETPL-LYSLVFGGVVNDATSVVFNALQSFDL 200
 DB 195 yxtxhffgailstqptvtaifnelhadvlyalifgesvindhavakxssivayqp 254
 QY 201 THLNHE-----AAPHLLGNFLYLLSTLGAAYGLISAVYIKKLYFGRHSTD-REVALM 254
 DB 255 aglnthafdaaaffkvgifgifsftngavtgvvtaxvtxk--ftkxhxpilletalf 312
 QY 255 MLMAYLSYMLAELFDLSILTVFCGVVMSHYTHWNVTSSRTTKHTTATLSFLAETFI 314
 DB 313 flmswtflaeacgfgtgvavlfctqahylynnlsvesrtrkqlfevhflaenfi 372
 QY 315 FLYVGMALDIDKWRVSVDPTGTSIAVSSILMG---LVMVGRAAFVFPISLSIAKKNQ 371
 DB 373 fsynglalftfk-----hvfspifilgafvaiflgraahtyplsfllnlgrrh- 421
 QY 372 SEKINFNQVVIWWSGLMRGAVSMALAYNKFTRAGHTDVRGNAMITSTITVCLFSTVVF 431
 DB 422 --kigwnfghmmfsgl-rgamafalair-----dtaeyarqmmfttlllvffvwi 472
 QY 432 GMLTKPLISYLLPHQNAATSMLSDDNTPKSIHPLLDQDSF 472
 DB 473 ggggttplmswl-----nirvgvdpdxpdp-----pxxdstaf 504

RESULT 14

AA44016

ID AA44016 standard; Protein; 569 AA.

XX

AC AA44016;

XX

DT 18-JAN-2000 (first entry)

XX

DE Schizosaccharomyces pombe Na/H antiport transporter.

XX

KW Sodium; proton; antiport; transporter; salt tolerance; salt management;
 KW transgenic plant; survival; soil; farming; accumulation; irrigation;
 KW crop.

XX

OS Schizosaccharomyces pombe.

XX

PN WO9947679-A2.

XX

PD 23-SEP-1999.

XX

PF 18-MAR-1999; 99WO-CA00219.

XX

PR 18-MAR-1998; 98US-0078474.
 PR 15-JAN-1999; 99US-0116111.
 XX (BLUM/) BLUMWALD E.
 PA (APSE/) APSE M.
 PA (SNED/) SNEDDEN W.
 PA (AHAR/) AHARON G.
 XX
 PI Blumwald E, Apse M, Snedden W, Aharon G;
 XX
 XX WPI; 1999-571840/48.

XX Nucleic acid molecules encoding sodium/proton transport polypeptides,
 PT useful in genetic engineering salt tolerance in crop plants -
 XX Disclosure; Fig 8A; 93pp; English.

XX The invention relates to an isolated nucleic acid molecule encoding
 CC a plant Na/H antiport (PNHX) transporter polypeptide, or a fragment
 CC and capable of increasing salt tolerance in a cell. This sequence
 CC corresponds to an Na/H antiport transporter from the fission yeast
 CC Schizosaccharomyces pombe. The Na/H transporter polypeptides provide
 CC a means of intracellular salt management, particularly in plants. The
 CC sequences are useful for producing transgenic plants that are capable
 CC of surviving in soil with high salt levels that would normally inhibit
 CC growth of the crop species. This would be useful in farming land in areas
 CC that are generally considered unproductive through salt accumulation and
 CC poor irrigation, e.g. in India, Australia, and prairies in USA or Canada.
 CC Commercial crops, such as potatoes, tomatoes, brassica, cotton, wheat,
 CC sunflower, strawberries, spinach, lettuce, rice, soybean, corn, wheat,
 CC rye, barley, atriplex, sorghum, alfalfa, salicornia and others would
 CC benefit from increased salt tolerance.

XX Sequence 569 AA;

Query Match 19.7%; Score 544; DB 20; Length 569;

Best Local Similarity 31.7%; Pred. No. 1.5e-48;

Matches 138; Conservative 103; Mismatches 157; Indels 38; Gaps 15;

QY 22 ALNLFVALLCIVLGHLEEN--RMNSESITALLIGLGTGVITLLISKKGK--SSHLVFS 78
 DB 36 alfillvlllgalltgyyqskkairahetvisvfgvmvvgllirvspglliqmvsfh 94
 QY 79 EDLFFIYLLPPIIFNAGVQKKQFFRNFVTIMLFGAVGTIISCTIISLGVTFQF----- 132
 DB 95 styfinvllppilnsgyelhgsnfrnigtlttfaqtffisa--vtlgvlyifsf 152
 QY 133 FKLDIGTFDLGDYLAIGALFAATDSVCTQLVNLQDET-PLLYSLVFGGVVNDATSVVV 191
 DB 153 fenlsm-tf--vealsmgatlsatdptvtaifnsykvqdqklytiifgesindhavalm 209
 QY 192 FNAIQSPDLTHLNEHAFLHLLGNFLYLLSTLGAATGLISAYVKKLYFGRHSTDREV 251
 DB 210 fetlqqfggtllhftfsgigifitffisilgvisglitalllkyslyrryps-ies 268
 QY 252 ALMMLMAYLSYMLAELFDLSILTVFCGVVMSHYTHWNVTSSRTTKHTTATLSFLAE 311
 DB 269 ciillmaysyffsngchmsgvslfcgtikhyaffmsykalstkyvfrvqla 328
 QY 312 TFIPLYVGMALD-IDDKWRVSVDPTGTSIAVSSILMGLVMVGRAAFVFPISLSIAKKN- 369
 DB 329 nfiiflygmslftqvdlvykpfifilattvavta-----srymavfplsnllakfhrq 380
 QY 370 ---NOSEKINFNQVVIWWSGLMRGAVSMALAYNKFTRAGHTDVRGNAMITSTITVCLF 426
 DB 381 rngnlidhypsqqmllfwagl-rgavgvla-----agfegenaqlrat-tlvvvv 432
 QY 427 STVVFGLTKPLISYL 442
 DB 433 tliifggtarmleil 448

Search completed: November 17, 2001, 13:27:37
Job time: 60 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 18, 2001, 09:52:16 ; Search time 155.97 Seconds
(without alignments)
8768.159 Million cell updates/sec

Title: US-09-271-584a-1

Perfect score: 2178
Sequence: 1 cctctgtttcttctctg.....aaaaaaaaaaaaaaaaaaaaa 2178

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_0601.*
1: /SID88/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2178	100.0	2178	20	Arabidopsis thalia
2	2143.4	98.4	2284	20	Arabidopsis thalia
3	832.8	38.2	1668	21	Attliplex gmelini N
4	821.2	37.7	2553	22	Gene regulating th
5	813	37.3	2423	22	Gene regulating th
6	774.8	35.6	2237	22	Gene regulating th
7	768.8	35.3	2330	21	cDNA encoding rice
8	754.4	34.6	2361	22	Gene regulating th
9	523.8	24.0	1788	20	Arabidopsis thalia
10	252.6	11.6	330	20	Arabidopsis Na/H a
11	226.4	10.4	378	20	Rice Na/H antiport

12	141.8	6.5	1824	21	AAC95372	Cat flea HMT Na/H
13	141.8	6.5	1824	21	AAC95373	Cat flea HMT Na/H
14	141.8	6.5	2080	21	AAC95370	Cat flea HMT Na/H
15	141.8	6.5	2080	21	AAC95371	Cat flea HMT Na/H
16	118.2	5.4	936	22	AAF58252	Oligonucleotide D1
17	118.2	5.4	936	22	AAF58253	Oligonucleotide D1
18	118.2	5.4	936	22	AAF58257	Oligonucleotide D1
19	118.2	5.4	936	22	AAF58259	Oligonucleotide D2
20	118.2	5.4	936	22	AAF58262	Oligonucleotide D2
21	118.2	5.4	936	22	AAF58255	Oligonucleotide D1
22	116.4	5.3	936	22	AAF58252	Oligonucleotide D1
23	116.4	5.3	936	22	AAF58254	Oligonucleotide D1
24	116.4	5.3	936	22	AAF58257	Oligonucleotide D1
25	116.4	5.3	936	22	AAF58259	Oligonucleotide D2
26	116.4	5.3	936	22	AAF58262	Oligonucleotide D1
27	116.4	5.3	936	22	AAF58255	Oligonucleotide D1
28	91.2	4.2	1581	22	AAF97895	Human secreted pro
29	82.2	3.8	1683	20	AAZ22602	Arabidopsis thalia
30	64.4	3.0	1354	22	AAF97931	Human secreted pro
31	64.4	3.0	1688	22	AAF97930	Human secreted pro
32	62.6	2.9	244	22	AAF58238	Oligonucleotide D1
33	60.2	2.8	714	20	AAZ22593	Arabidopsis thalia
34	60.2	2.8	2122	20	AAZ22603	Arabidopsis thalia
35	58.6	2.7	268	20	AAZ22607	Rice Na/H antiport
36	57.2	2.6	244	22	AAF58238	Oligonucleotide D1
37	55.8	2.6	1244	20	AAZ97674	Extended human sec
38	55.4	2.5	1347	20	AAZ97773	Extended human sec
39	53.8	2.5	1291	21	AAZ53975	TRIP-AP53 tumour s
40	53.8	2.5	2007	19	AAV40754	C. felis esterase,
41	53.8	2.5	2007	19	AAV40755	C. felis esterase,
42	52.2	2.4	5852	12	AAQ11710	Dictyostellium plas
43	52	2.4	263	21	AAA44056	Human secreted exp
44	48.8	2.2	682	21	AAA44930	Human secreted exp
45	48.8	2.2	1169	20	AAV99983	3' untranslated re

ALIGNMENTS

RESULT 1
AAZ22591
ID AAZ22591 standard; cDNA; 2178 BP.
XX
AC AAZ22591;
XX
DT 18-JAN-2000 (first entry)
XX
DE Arabidopsis thaliana Na/H transporter gene AtNHX1.
XX
KW Sodium; proton; antiport; transporter; salt tolerance; salt management;
KW transgenic plant; survival; soil; farming; accumulation; irrigation;
KW crop: ss.

OS Arabidopsis thaliana.
XX
XX
PN WO9947679-A2.
XX
PD 23-SEP-1999.
XX
PF 18-MAR-1999; 99WO-CA00219.
XX
PR 18-MAR-1998; 98US-0078474.
PR 15-JAN-1999; 99US-0116111.
XX
PA (BLUM/) BLUMWALD E.
PA (APSE/) APSE M.
PA (SNED/) SNEDDEN W.
PA (AHAR/) AHARON G.
XX
PI Blumwald E, Apse M, Snedden W, Aharon G;
XX
DR WPI: 1999-571840/48.
DR P-PSDB; AAY40901.

XX Nucleic acid molecules encoding sodium/proton transport polypeptides,
PT useful in genetic engineering salt tolerance in crop plants -
XX
PS
XX
XX Claim 4; Fig 1A; 93bp; English.
XX
XX The invention relates to an isolated nucleic acid molecule encoding a
CC plant Na/H antiport (PNHX) transporter polypeptide, or a fragment and
CC capable of increasing the salt tolerance in a cell. This sequence corresponds
CC to the gene encoding the ANHX1 transporter from Arabidopsis thaliana.
CC The Na/H transporter polypeptides provide a means of intracellular salt
CC management, particularly in plants. The sequences are useful for
CC producing transgenic plants that are capable of surviving in soil with
CC high salt levels that would normally inhibit growth of the crop species.
CC This would be useful in farming land in areas that are generally
CC considered unproductive through salt accumulation and poor irrigation,
CC e.g. in India, Australia, and prairies in USA or Canada. Commercial
CC crops, such as potatoes, tomatoes, brassica, cotton, sunflower,
CC strawberries, spinach, lettuce, rice, soybean, corn, wheat, rye, barley
CC atriplex, sorghum, alfalfa, salicornia and others would benefit from
CC increased salt tolerance.
XX
XX Sequence 2178 BP: 576 A; 413 C; 472 G; 717 T; 0 other;
SQ

Query Match 100.0%; Score 2178; DB 20; Length 2178;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2178; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy 1 cctctctgttctctctgtatgacaaagaaagaaagaaatcaggttttagctttcgca 60
Db 1 cctctctgttctctctgtatgacaaagaaagaaagaaatcaggttttagctttcgca 60
Qy 61 agcttcacaaatttgaatttgaattctctgggctcttttgaatcagactgaagatat 120
Db 61 agcttcacaaatttgaatttgaattctctgggctcttttgaatcagactgaagatat 120
Qy 121 ttgattaccagaagttgttcaagaaatgtttcagtgacagacagcagaaagataaaag 180
Db 121 ttgattaccagaagttgttcaagaaatgtttcagtgacagcagcagaaagataaaag 180
Qy 181 agacttttttccagattttgctgataccaaaatcgaatgttgctcatgttttgat 240
Db 181 agacttttttccagattttgctgataccaaaatcgaatgttgctcatgttttgat 240
Qy 241 caaatctggaagaggagttgttggatctagaagaagatacaaatgttgatctctca 300
Db 241 caaatctggaagaggagttgttggatctagaagaagatacaaatgttgatctctca 300
Qy 301 gtgtcgaaactccttcgttatcgacatctgatcacgcttctgtggttcggtgaatctc 360
Db 301 gtgtcgaaactcgttcgttatcgacatctgatcacgcttctgtggttcggtgaatctc 360
Qy 361 tttgtgcaactcttgtgctgttattgttcttggtoatcttttggaaagaaatagatgg 420
Db 361 tttgtgcaactcttgtgctgttattgttcttggtoatcttttggaaagaaatagatgg 420
Qy 421 atgaacgaatccatcaacgcttcttgaattgggtcagtcagctggtgttaccattttgtg 480
Db 421 atgaacgaatccatcaacgcttcttgaattgggtcagtcagctggtgttaccattttgtg 480
Qy 481 attagtaagaaagaaagctgcactctctctcgtctttagtgaagatcttttctcatatat 540
Db 481 attagtaagaaagaaagctgcactctctctcgtctttagtgaagatcttttctcatatat 540
Qy 541 cttttgcaacccattatatcaatgcaggggtttcaagtaaaaaagacagttttccgc 600
Db 541 cttttgcaacccattatatcaagtcaggggtttcaagtaaaaaagacagttttccgc 600
Qy 601 aatttcgtgactattatgcttttgggtgctgttgggactattattcttgcacaatcata 660
Db 601 aatttcgtgactattatgcttttgggtgctgttgggactattattcttgcacaatcata 660

Db 1741 agtatacgtgctcttgacacgcccactcgaacgctgcattactacggacaaattt 1800
Qy 1801 gatgactccctcagcagaccgctcttgaggctgctggttgtaacctttgtccaggt 1860
Db 1801 gatgactccctcagcagaccgctcttgaggctgctggttgtaacctttgtccaggt 1860
Qy 1861 tctccactgagagaacccctcctgatcttagtaagccttgagggttaacgtggaagaaa 1920
Db 1861 tctccactgagagaacccctcctgatcttagtaagccttgagggttaacgtggaagaaa 1920
Qy 1921 gctttgatttttttgtagaagaaggttgattcaaatattgcttttgtaattatcca 1980
Db 1921 gctttgatttttttgtagaagaaggttgattcaaatattgcttttgtaattatcca 1980
Qy 1981 ttgtgaataattgttgtaggacagaaatcgtctcctaaacgttttgagagcagaagcaaa 2040
Db 1981 ttgtgaataattgttgtaggagcagaaatcgtctcctaaacgttttgagagcagaagcaaa 2040
Qy 2041 acatggcaactttgaagtggttgattgatgtatgtaattatattcatattttgtgtg 2100
Db 2041 acatggcaactttgaagtggttgattgatgtatgtaattatattcatattttgtgtg 2100
Qy 2101 taacacaaactacacattgtttgtattgttttgattggtttttgttcgaaaaaaaaa 2160
Db 2101 taacacaaactacacattgtttgtattgttttgattggtttttgttcgaaaaaaaaa 2160
Qy 2161 aaaaaaaaaaaaaaaaaa 2178
Db 2161 aaaaaaaaaaaaaaaaaa 2178

RESULT 2
AAZ22595
ID AAZ22595 standard; cDNA; 2284 BP.
AC AAZ22595;
XX
XX
DT 18-JAN-2000 (first entry)
DE Arabidopsis thaliana Na/H transporter gene.
XX
XX Sodium; proton; antiport; transporter; salt tolerance; salt management;
KW transgenic plant; survival; soil; farming; accumulation; irrigation;
KW crop; ss.
XX
OS Arabidopsis thaliana.
XX
XX W09947679-A2.
XX
PD 23-SEP-1999.
XX
PF 18-MAR-1999; 99WO-CA00219.
XX
PR 18-MAR-1998; 98US-0078474.
PR 15-JAN-1999; 99US-0116111.
XX
XX (BLUM/) BLUMWALD E.
PA (APSE/) APSE M.
PA (SNED/) SNEDDEN W.
PA (AHAR/) AHARON G.
XX
PI Blumwald E, Apse M, Snedden W, Aharon G;
XX
DR WPI; 1999-571840/48.
DR P-PSDB; AAY40905.
XX
XX Nucleic acid molecules encoding sodium/proton transport polypeptides,
PT useful in genetic engineering salt tolerance in crop plants -
XX
XX Claim 48; Fig 5A-B; 93pp; English.
PS
XX The invention relates to an isolated nucleic acid molecule encoding

CC a plant Na/H antiport (PNHX) transporter polypeptide, or a fragment
CC and capable of increasing salt tolerance in a cell. This sequence
CC corresponds to the gene encoding a transporter from Arabidopsis thaliana.
CC The Na/H transporter polypeptides provide a means of intracellular salt
CC management, particularly in plants. The sequences are useful for
CC producing transgenic plants that are capable of surviving in soil with
CC high salt levels that would normally inhibit growth of the crop species.
CC This would be useful in farming land in areas that are generally
CC considered unproductive through salt accumulation and poor irrigation,
CC e.g. in India, Australia, and prairies in USA or Canada. Commercial
CC crops, such as potatoes, tomatoes, brassica, cotton, sunflower,
CC strawberries, spinach, lettuce, rice, soybean, corn, wheat, rye, barley,
CC atriplex, sorghum, alfalfa, salicornia and others would benefit from
CC increased salt tolerance.
XX
SQ Sequence 2284 BP; 589 A; 444 C; 485 G; 766 T; 0 other;

Query Match 98.4%; Score 2143.4; DB 20; Length 2284;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2166; Conservative 0; Mismatches 1; Indels 2; Gaps 2;
Qy 1 cctctgtttcgtctcgtagacgaagaagaagaatctcaggttttagctttcga 60
Db 116 cctctgtttcgtctcgtagacgaagaagaagaatctcaggttttagctttcga 175
Qy 61 agcttccaaaaatttgaaatttgatctctctggtctcttttgtaaatcagactgaagatat 120
Db 176 agcttccaaaaatttgaaatttgatctctctggtctcttttgtaaatcagactgaagatat 235
Qy 121 ttgattaccacagaagttgttcaaggaaatggtttcagtgacagcagcagcgaagataaaag 180
Db 236 ttgattaccacagaagttgttcaaggaaatggtttcagtgacagcagcagcgaagataaaag 295
Qy 181 agactttttttccagattttgctgacccaaatctgaatagttgttcattgtcttgat 240
Db 296 agactttttttccagattttgctgacccaaatctgaatagttgttcattgtcttgat 355
Qy 241 caaatctggaagagaagttgttgatctagaaagataaacaatgttggtattctcta 300
Db 356 caaatctggaagagaagttgttgatctagaaagataaacaatgttggtattctcta 415
Qy 301 gtgtcgaactgctctcgttctatcgacatctgacacgcttctgtgttgcttgaaatctc 360
Db 416 gtgtcgaactgctctcgttctatcgacatctgacacgcttctgtgttgcttgaaatctc 475
Qy 361 ttgtgtgactttttgtgtgtgtattgttctgttgcttcattcttttggaagagaatagtg 420
Db 476 ttgtgtgactttttgtgtgtgtattgttctgttgcttcattcttttggaagagaatagtg 535
Qy 421 atgaacgaatccatcacccgcttgttgattgggtgactgagcactggtttaccatttgttg 480
Db 536 atgaacgaatccatcacccgcttgttgattgggtgactgagcactggtttaccatttgttg 595
Qy 481 attagtaagggaaaaagctcgcatcttctcgtcttttagtgaagatctttcttcacatat 540
Db 596 attagtaagggaaaaagctcgcatcttctcgtcttttagtgaagatctttcttcacatat 655
Qy 541 cttttgcccaccattattatccaatgcagggtttccaagttaaaaaaagcagtttttccgc 600
Db 656 cttttgcccaccattattatccaatgcagggtttccaagttaaaaaaagcagtttttccgc 715
Qy 601 aatttcgtgactattatgctttttgtgctgttggaactattattcttcgcaaatcata 660
Db 716 aatttcgtgactattatgctttttgtgctgttggaactattattcttcgcaaatcata 775
Qy 661 tctctagggttaacacagttctttaaagaagttggacattggaaacctttgacttgggtgat 720
Db 776 tctctagggttaacacagttctttaaagaagttggacattggaaacctttgacttgggtgat 835
Qy 721 tatcttgctattggtgccatatttgcgcgaacagatctcagtagtacactgcaggttctg 780
Db 836 tatcttgctattggtgccatatttgcgcgaacagatctcagtagtacactgcaggttctg 895

QY 781 aatcaagcagagacacotttgcctttacagttctgtattcggaggggtgtgtgaatgat 840
Db 896 aatcaagcagagacacotttgcctttacagttctgtattcggaggggtgtgtgaatgat 955
QY 841 gcaacgtcaagttgtgtctcaacgcgattcagagctttgatctcaactcaactaaacac 900
Db 956 gcaacgtcaagttgtgtctcaacgcgattcagagctttgatctcaactcaactaaacac 1015
QY 901 gaagctgtcttttcattctcttcttggaacctcttctgtatttcttctcctaagtaaccttgcct 960
Db 1016 gaagctgtcttttcattctcttcttggaacctcttctgtatttcttctcctaagtaaccttgcct 1075
QY 961 ggtgtcgaacccggtctgataagtcgctgatttattcaagaagcctataacttttggaagccac 1020
Db 1076 ggtgtcgaacccggtctgataagtcgctgatttattcaagaagcctataacttttggaagccac 1135
QY 1021 tcaactgaccgagaggttgcctcttatgatcttattggtcgtatcttcttcttatatgcttgcct 1080
Db 1136 tcaactgaccgagaggttgcctcttatgatcttattggtcgtatcttcttcttatatgcttgcct 1195
QY 1081 gagcttttgcacttgagcgggtatctcactgctgttttctctgtgtattgtattgtatgcctat 1140
Db 1196 gagcttttgcacttgagcgggtatctcactgctgttttctctgtgtattgtattgtatgcctat 1255
QY 1141 tacacatgccaatgttaacggagagctcaagaatacaacaagaacataccttttgcacct 1200
Db 1256 tacacatgccaatgttaacggagagctcaagaatacaacaagaacataccttttgcacct 1315
QY 1201 ttgtcaattcttcgagagacatttatttctgtatgttggaatggatgccttggacatt 1260
Db 1316 ttgtcaattcttcgagagacatttatttctgtatgttggaatggatgccttggacatt 1375
QY 1261 gacaagtggagatccgtgagtgacacacgggaacatcgatcgagtgagctcaatcccta 1320
Db 1376 gacaagtggagatccgtgagtgacacacgggaacatcgatcgagtgagctcaatcccta 1435
QY 1321 atgggtctggtcattggttggaagagcagcgttctcgtcttccgtttatcgtttcttatctaac 1380
Db 1436 atgggtctggtcattggttggaagagcagcgttctcgtcttccgtttatcgtttcttatctaac 1495
QY 1381 ttagccaagaagaatacaacgcgagaaaaatcaactttaacatgcaggttgcatttgggtgg 1440
Db 1496 ttagccaagaagaatacaacgcgagaaaaatcaactttaacatgcaggttgcatttgggtgg 1555
QY 1441 tctggctcattgagaggtgctgtatctatgctcttgatcaacaagaatttacaagggcc 1500
Db 1556 tctggctcattgagaggtgctgtatctatgctcttgatcaacaagaatttacaagggcc 1615
QY 1501 gggcacacagatgtacgcgggaatgcaatcatatgatacagatcacataactgtctgtctt 1560
Db 1616 gggcacacagatgtacgcgggaatgcaatcatatgatacagatcacataactgtctgtctt 1675
QY 1561 tttagcacagtggtgttggatgtcgtacaaacacactcaatgaactacattatccgcac 1620
Db 1676 tttagcacagtggtgttggatgtcgtacaaacacactcaatgaactacattatccgcac 1735
QY 1621 cagaacgcccacacagatgttattctgatgacaacaccccacaaatccatacatatccct 1680
Db 1736 cagaacgcccacacagatgttattctgatgacaacaccccacaaatccatacatatccct 1795
QY 1681 ttgttgaccaagaactcgttctatgagcctttaggagccttaagggaacacaaatgctcgtggcctgac 1740
Db 1796 ttgttgaccaagaactcgttctatgagcctttaggagcctttagggaacacaaatgctcgtggcctgac 1855
QY 1741 agtatcgtgtgcttcttgacacggccactc-gaacgtgcattact-actggagacat 1798
Db 1856 agtatcgtgtgcttcttgacacggccactcggaaacccggaacccgtaactaactaggagacat 1915
QY 1799 ttgatgactcttctcatgcgaccgctctttggaggtcgtgtggttctgtacacctttgttccag 1858
Db 1916 ttgatgactcttctcatgcgaccgctctttggaggtcgtgtggttctgtacacctttgttccag 1975

QY 1859 gttctccaactgagagaaacccctcctgatcttagtaagcgttgagggttaacgtggaagaa 1918
Db 1976 gttctccaactgagagaaacccctcctgatcttagtaagcgttgagggttaacgtggaagaa 2035
QY 1919 aagctttgatttttttttgtagaaaagggtgattcaaatatgcttttgcgttaaatattc 1978
Db 2036 aagctttgatttttttttgtagaaaagggtgattcaaatatgcttttgcgttaaatattc 2095
QY 1979 catttgtatatttttttgtagagacagaaatctgtcctaactgttttgagagcagaaagca 2038
Db 2096 catttgtatatttttttgtagagacagaaatctgtcctaactgttttgagagcagaaagca 2155
QY 2039 aaacatggcaacctttgaagtgtttgatgtatgtatgaattatattcatatttgtttgt 2098
Db 2156 aaacatggcaacctttgaagtgtttgatgtatgtatgaattatattcatatttgtttgt 2215
QY 2099 tgaacacacaaactacacatttgtttatgttttgaattgttttgcgttcgaaaaaaaa 2158
Db 2216 tgaacacacaaactacacatttgtttatgttttgaattgttttgcgttcgaaaaaaaa 2275
QY 2159 aaaaaaaa 2167
Db 2276 aaaaaaaa 2284
RESULT 3
AAAT72926
ID AAA72926 standard; DNA; 1668 BP.
XX AC AAA72926;
XX XX
DT 23-NOV-2000 (first entry)
XX DE Atliplex gmelini Na+ and H+ antiporter protein encoding DNA.
XX KW Atliplex gmelini; Na plus and H plus antiporter protein;
KW Na+ and H+ antiporter protein; transformed plant;
XX KW high salt tolerance; ds.
XX OS Atliplex gmelini.
XX PN JP2000157287-A.
XX PD 13-JUN-2000.
XX PF 16-SEP-1999; 99JP-0261606.
XX PR 24-SEP-1998; 98JP-0269504.
XX PA (SHOK-) SHOKUBUTSU KOGAKU KK.
XX DR WPI; 2000-468209/41.
DR P-PSDB; AAB12786.
XX PT An Na+ and H+ plus antiporter protein and a gene encoding it
XX PS Claim 1; Page 9-10; 16pp; Japanese.
XX CC The present sequence encodes an Na+ and H+ antiporter protein isolated
CC from Atliplex gmelini. The Na+ and H+ antiporter protein and gene
CC encoding it are useful for the preparation of transformed plants with
CC high salt tolerance, e.g. for growth in arid land.
XX SQ Sequence 1668 BP; 397 A; 342 C; 376 G; 553 T; 0 other;

Query Match 38.2%; Score 832.8; DB 21; Length 1668;
Best Local Similarity 71.9%; Pred. No. 1.1e-176;
Matches 1146; Conservative 0; Mismatches 412; Indels 36; Gaps 3;
QY 323 cgacatctgatacagctctctgtgtgcgttggaatctcttttgcactcttctgtgctt 382
Db 50 ccaactctgatacagctctctgtgtgcgttggaactgtttgtgcactgttatgtggt 109

CC new breeds of colourful plants for cut flowers, particularly applicable
CC in horticulture.
XX Sequence 2553 BP; 666 A; 487 C; 535 G; 865 T; 0 other;

Query Match		37.7%;	Score 821.2;	DB 22;	Length 2553;
Best Local Similarity		70.3%;	Pred. No. 4.7e-174;		
Matches 1115;		Conservative 0;	Mismatches 468;	Indels 3;	Gaps 1;
Qy	289	ttggattctctagtcgaaacagtcgcttcgttatcgacatctgatcagcgtctctgtggtt	348		
Db	543	tttggactctctggaagatgaacaacttaacaactctgatcatcaatcagtggt	602		
Qy	349	gggtgaactctcttggctgcaacttctgtgtatgttcttctgctcatcttttgaa	408		
Db	603	tcggtaaacctgttctgacactatttgcgctgtatcttgctgctgcttattctgag	662		
Qy	409	gagaatagatgatgaacgaatccatcacccgcttctgttattggctaggcactgggtt	468		
Db	663	gaaaacagatgatgaatgagtcacactgacactgcttctgtgattggtgactggagtc	722		
Qy	469	acatttttctgattagtaaaagaaagctgcgactctctctgttcttctgtaggaagatctt	528		
Db	723	atctctactaataagtggaggaagaaactcacatatatttagttgttcagcgaagatctt	782		
Qy	529	ttctctatatcttttgcacccattatccaatgcaggggtttccaagttaaaaaagaag	588		
Db	783	ttctctattacccttctccacgatcatttttaagtctgggttccaggtgaaaaagaaa	842		
Qy	589	cagttttccgcaattctgactattatgtcttttggctgtgctgtggactattattct	648		
Db	843	tcatttccgcaattctgactatcatctgtcttcttgggcagttggcacttggatctg	902		
Qy	649	tgacaatatatctctaggttaaacacagttctttaagaagtggacattggaaccttt	708		
Db	903	ttcatktatatacagcgggtgctatctggcatcttccaagaataaggatattggacacctt	962		
Qy	709	gacttgggtgattatctgtctattgttgcacatatttctgccaacagattcagtagtaca	768		
Db	963	gaaatggagattaaccttgcgaatggagcaactcttctgccaacagattctgtatgcacc	1022		
Qy	769	ctgcaggttctgaatacgaacagacacaccttctgtctttaaagctctgtatccgagaggt	828		
Db	1023	ttacaagtgcttaatacagaagaacacccgttattgtacagctcagttgttgggaaggt	1082		
Qy	829	gttgtgaatgatcaacgtcagttgttctccaacgcgattcagagcttctgtatctcact	888		
Db	1083	gttgtgaatgatccacatctgtatgtctgttccaatgtctgccaacactttgacttatct	1142		
Qy	889	cacctaaaccacgaagctgctttctcatctcttcttggaactcttctgtattgtttctccta	948		
Db	1143	catatcagcacaggcaagctctgcgaatttaattggaaactttctatactgtttgctcgc	1202		
Qy	949	agtaacttgttggtgctgcaacccgtctgataagtgcatgtattatcaagaagctatcac	1008		
Db	1203	agcaacttctaggggtgtgtgttggccactaaggcttcttaataattgaagaactctac	1262		
Qy	1009	tttggaaaggcactcaactgaccgagaggttgcctcttatgatgcttattggcgtatcttct	1068		
Db	1263	tttggaaaggcactgactgatctgaggttgcgtataatgatactcatgctgctacatcata	1322		
Qy	1069	tatatgtctgagctttctgcactgagcgttatctcactgtgtttttctctgtggtatt	1128		
Db	1323	tacatgtctggaattatctattttaaagtggaactcctcactgtgtttctctgtgggact	1382		
Qy	1129	gtatgtcccatcacatagccacaaatgtaacggagagctcaagaataacaacaagaact	1188		
Db	1383	gtgatgtctcactatactctggcataaagtgtactgagagctcaagagtcactaccaagcac	1442		
Qy	1189	acctttgcaactttgtcatcttctgcggagacatttattttctgtatgttggaaaggat	1248		
Db	1443	acgtttgctacattcatctatttctgctgaaataatccatattctcttattgttggatggat	1502		

Qy	1249	gccttggacattgacaaagtgagatccgtgagtgcacacacccgggaacatcgatcgagtg	1308
Db	1503	gccttggacattgagaagtggaagtgttaagcgacagcccggaacatcaatgaagtc	1562
Qy	1309	agctcaatcctaataatgggtctgtcatcgttggaaagagcagcgttcttctccgttatcg	1368
Db	1563	agctcaatctgtcaggtctctgttgggttggaaagggagccttgttttccctctgtca	1622
Qy	1369	tttctatcaactgaccgaagaatcaaacgagaaatacaacttttaacatgcaggtt	1428
Db	1623	ttcttgcacacttgaccgaagaaatactcgaggacaaagattagctttaaccagcaggtt	1682
Qy	1429	gtgatttgggtctgtctcattgagaggtgtctgtatctatctatggctcttgcatacaacaag	1488
Db	1683	acaatatgggtgggtggttgcgaggtgtctgttctatggccttgccttataatcag	1742
Qy	1489	tttacaaggccgggcacacagatgtacgcgggaatgcaatcatcatcacgagtagcagata	1548
Db	1743	tttaccaggggaggtcatactcagttcacgtgcacaaatgcataatgatcagagtagctatc	1802
Qy	1549	actgtctgtctttttagcacagtggttgggttatgtctgaccaaaacc---actcataaagc	1605
Db	1803	actgtgtccttttcagcacagtggttatttgggttgcagacaaacactttaattctatta	1862
Qy	1606	tactattacgcacacgaacgcacacacagcagatgttattctatgcatacaacaccccaaaa	1665
Db	1863	ttgtaccttcacaaaacacacttgatcagaatgatctctctgaaaccgtagctcccaaaa	1922
Qy	1666	tcatacatatcctcttggccaagactcgttcatgtgaccttcagggaacccaact	1725
Db	1923	tcctcatgtgcacacttctgacagcacacaagactcagaagctgatctggtgcgacat	1982
Qy	1726	gtgctcgggctgacagatatacgttggcttcttgcacacggcccaactcgaaaccgtgacatc	1785
Db	1983	gtaccccgctcccaacagtttgcgagtgctctctgtcaaccccatctcacacggtacattac	2042
Qy	1786	tactggagacaatttgatgactctctcatcgacccgcttcttggaggtcgtgttggta	1845
Db	2043	tactgggaaaaatttgacaatgcattcagctcgtctgttctcggtagcagaggttttgta	2102
Qy	1846	ccctttgttccaggttctccaactga	1871
Db	2103	ccctttgttccaggtacacactatga	2128

RESULT 5
AAF75764
ID AAF75764 standard; DNA; 2423 BP.
XX
AC AAF75764;
XX
DT 14-MAY-2001 (first entry)
XX
DE Gene regulating the pH of vacuoles.
XX
KW Vacuole pH regulation; flower colour; ds.
XX
OS Petunia hybrida.
XX
PN WO200114560-A1.
XX
PD 01-MAR-2001.
XX
PF 24-AUG-2000; 2000WO-JP05722.
XX
PR 24-AUG-1999; 99JP-0236800.
XX
PA (SUNR) SUNTORY LTD.
XX
PI Iida S, Tanaka S, Inagaki Y;
XX
DR WPI; 2001-191648/19.

Db 941 gctcaagaatcaactccaggcatgtatttccaatgttgcctttattgcggaacattca 1000
Qy 1226 tttcttgatgttggaatgagtccttggaacattgacaagtggagatccgtgagtgaca 1285
Db 1001 tattctgtatgttggaacagatgctcttgattttacaaagtggagacaaagcagctta 1060
Qy 1286 caccgggaacatcgatcgacgtgagctcaatcctaattggtctggtcatgtgttggaaag 1345
Db 1061 gctttggggtactctgggtgtcccggtgtcataacccgcatgtattgcttggacgag 1120
Qy 1346 cagcgttctgttccgttactgtttctatctaaacttagccaaagaagaa-----tcaaa 1399
Db 1121 cagcatttcttccactctcggtctttaaaaaattcatgaacaggcacaactgaaagaa 1180
Qy 1400 gcgagaaaatacaatttaacatgcaggtgtgtgatttgggtggtctctcatgagaggtg 1459
Db 1181 acgagctatcacatttaagaatcagtcagtgatcatttgggtggcaggtctaatgcgaggtg 1240
Qy 1460 ctgtatctatgctcttgcatatacaaaagtttacaaggccgggacacacagatgtacgg 1519
Db 1241 ctgtctcaatgtcctggcttccagcagttccacatactccggtgttacattggtactcg 1300
Qy 1520 ggaatgcaatcatgatcacgagtagcagataactgtctgtcttttttagcaagtggtgttg 1579
Db 1301 tgaatgtgccatgttcaccaaacaccactatcgtgttctctttactacactggtctttg 1360
Qy 1580 gtatctgacaaaacacatcataagtcactattaccgcacca---gaaagccacacaga 1636
Db 1361 gttctctacaaaacactgtgaaattatctctctcccaagatgcgaagtcacacacagg 1420
Qy 1637 gcatgttatctgatgacacaccccaaaatccatacatcctcttgttggaccaagact 1696
Db 1421 gaaatagaggttaaacgcactgagccaggttctccgaaagaagatgcacacttctcttc 1480
Qy 1697 cgttcattgagccttcagggaacacaaatgtgcctcggcct---gacagttacgtggct 1753
Db 1481 ttctcttgacgagctcgtcttcacacaaacttcaatagagctagagatagtattcccttc 1540
Qy 1754 tcttgacacggccactcgaaacgttgactactactgagagacaatttgactccttca 1813
Db 1541 tgaatgaaacacctgttacacacatccaccgctactgagaaagtgtgacgacacataca 1600
Qy 1814 tgcgaccgctctttggaggtcgtgg 1638
Db 1601 tgaggcctatctcgttggaacctg 1625

RESULT 10
AAZ22611
ID AAZ22611 standard; DNA; 330 BP.
XX
AC AAZ22611;
XX
DT 18-JAN-2000 (first entry)
XX
DE Arabidopsis Na/H antiport transporter gene from EST clone r75860.
XX
KW Sodium; proton; antiport; transporter; salt tolerance; salt management;
KW transgenic plant; survival; soil; farming; accumulation; irrigation;
KW crop; ss.
XX
OS Arabidopsis thaliana.
XX
PN WO9947679-A2.
XX
PD 23-SEP-1999.
XX
PF 18-MAR-1999; 99NO-CA00219.
XX
PR 18-MAR-1998; 98US-0078474.
PR 15-JAN-1999; 99US-0116111.
XX
PA (BLUM/) BLUMWALD E.

PA (APSE/) APSE M.
PA (SNED/) SNEDDEN W.
XX (AHAR/) AHARON G.
PI Blumwald E, Apse M, Snedden W, Aharon G;
XX WPI; 1999-571840/48.
DR Nucleic acid molecules encoding sodium/proton transport polypeptides,
XX useful in genetic engineering salt tolerance in crop plants -
PT Disclosure: Fig 8H; 93pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule encoding
CC a plant Na/H antiport (PNHX) transporter polypeptide, or a fragment
CC and capable of increasing salt tolerance in a cell. This sequence
CC corresponds to an Na/H antiport transporter gene fragment from the
CC Arabidopsis expressed sequence tag (EST) library clone AAT75860. The
CC Na/H transporter polypeptides provide a means of intracellular salt
CC management, particularly in plants. The sequences are useful for
CC producing transgenic plants that are capable of surviving in soil
CC with high salt levels that would normally inhibit growth of the crop
CC species. This would be useful in farming land in areas that are
CC generally considered unproductive through salt accumulation and poor
CC irrigation, e.g. in India, Australia, and prairies in USA or Canada.
CC Commercial crops, such as potatoes, tomatoes, brassica, cotton, wheat,
CC sunflower, strawberries, spinach, lettuce, rice, soybean, corn, wheat,
CC rye, barley, atriplex, sorghum, alfalfa, salicornia and others would
CC benefit from increased salt tolerance.
XX Sequence 330 BP; 92 A; 68 C; 66 G; 91 T; 13 other;
SQ

Query Match 11.6%; Score 252.6; DB 20; Length 330;
Best Local Similarity 90.7%; Pred. No. 2.1e-47;
Matches 264; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
Qy 1357 ttccgttatcgtttctatctaaacttagccaagaagaatacaagcgagaaatcaacttt 1416
Db 1 ttccgttatcgtttctatctaaacttagccaagaagaatacaagcgagaaatcaacttt 60
Qy 1417 aacatgcaggttgatttgggtggtctctcatgagaggtgcttatctatggtcttt 1476
Db 61 aacatgcaggttgatttgggtggtctctcatgagaggtgcttatctatggtcttt 120
Qy 1477 gcatacaacaagtttacaaggccgggacacacagatgtacgcgggaatgcaatcatgac 1536
Db 121 gcatacaacaagtttacaaggccgggacacacagatgtacgnggggaatgcaatcatgac 180
Qy 1537 acgagtacgataactgtctgtctcttttttagcacagtgtgtttgtgtgctgaccaaacca 1596
Db 181 acngtgacntaactgtctgtnttttttagcacagtgtgtttgtgtgctgaccaaacca 240
Qy 1597 ctcaataagctacctattaccgcaccagaacgcgcacacagcgcatgttatct 1647
Db 241 ntcaataagctacctattaccgcaccagaacgcgcacacgctcatcaccnggcatgt 291

RESULT 11
AAZ22606
ID AAZ22606 standard; DNA; 378 BP.
XX
AC AAZ22606;
XX
DT 18-JAN-2000 (first entry)
XX
DE Rice Na/H antiport transporter gene from EST clone C91832.
XX
KW Sodium; proton; antiport; transporter; salt tolerance; salt management;
KW transgenic plant; survival; soil; farming; accumulation; irrigation;
KW crop; ds.
XX
OS Oryza sativa.

```
XX PN WO9947679-A2.
XX PD 23-SEP-1999.
XX PF 18-MAR-1999; 99WO-CA00219.
XX PR 18-MAR-1998; 98US-0078474.
XX PS 15-JAN-1999; 99US-0116111.
XX PA (BLUM/) BLUMWALD E.
XX PA (APSE/) APSE M.
XX PA (SNED/) SNEDDEN W.
XX PA (AHAR/) AHARON G.
XX PI Blumwald E, Apse M, Snedden W, Aharon G;
XX PF WPI; 1999-571840/48.
XX PT Nucleic acid molecules encoding sodium/proton transport polypeptides,
XX PS useful in genetic engineering salt tolerance in crop plants -
XX PS Disclosure; Fig 8C; 93pp; English.
XX CC The invention relates to an isolated nucleic acid molecule encoding
XX CC a plant Na/H antiport (PNHX) transporter polypeptide, or a fragment
XX CC and capable of increasing salt tolerance in a cell. This sequence
XX CC corresponds to an Na/H antiport transporter gene fragment from the
XX CC rice expressed sequence tag (EST) library clone AAC91832. The Na/H
XX CC transporter polypeptides provide a means of intracellular salt
XX CC management, particularly in plants. The sequences are useful for
XX CC producing transgenic plants that are capable of surviving in soil
XX CC with high salt levels that would normally inhibit growth of the crop
XX CC species. This would be useful in farming land in areas that are
XX CC generally considered unproductive through salt accumulation and poor
XX CC irrigation, e.g. in India, Australia, and prairies in USA or Canada.
XX CC Commercial crops, such as potatoes, tomatoes, brassica, cotton,
XX CC sunflower, strawberries, spinach, lettuce, rice, soybean, corn, wheat,
XX CC rye, barley, acirplex, sorghum, alfalfa, salicornia and others would
XX CC benefit from increased salt tolerance.
XX SQ Sequence 378 BP; 88 A; 74 C; 89 G; 121 T; 6 other;

Query Match 10.4%; Score 236.4; DB 20; Length 378;
Best Local Similarity 77.2%; Pred. No. 1.5e-41;
Matches 284; Conservative 0; Mismatches 82; Indels 2; Gaps 1;

QY 996 caagaagctatcttgggaagcactcaactgaccgagaggttgcccttatgactcttat 1055
DB 1 caagaagctatctatcttgggaagcattctactgacctgaggttgcccttatgactctat 60
QY 1056 ggcgtatcttttatatcttgcgtgagcttttcgacttgagcggtatcactctgtgtt 1115
DB 61 ggcttacctttcatatctgctgagctgtgtagatttgagcggtatcactctacogtatt 120
QY 1116 ttctgtgtattgtgattgctccattacacatggtgcacatgtaacgagagctcaagaat 1175
DB 121 ctctctgtgtattgtaattgtcacattacacttggcactacgtccacagagaggttcaagagt 180
QY 1176 aacacaaagacatacttgcgaactttgcaatttcgacttgagcggtatcactctgttata 1235
DB 181 tacaacaaagacgatttgcaactctctctctctctctctctctctctctctctctctctga 240
QY 1236 ttttggaaatggatgcttggaacattgacaagtggagatccgtgagtcacacacccgggaac 1295
DB 241 ttttgggaatggatgcttggaatggaatgaaatggagatnccagtgacagacctggnaa 300
QY 1296 atcgatcgacgtgagctcaatcct--aatgggtctgtctatggttggaagacagcgttc 1353
DB 301 atccattgggtgaagctcaatttctgtagggattggttctctgattgggaagngctgcttt 360
QY 1354 gctcttccc 1361
```

Db 361 gnaattccc 368

RESULT 12

AAC95372

ID AAC95372 standard; cDNA; 1824 BP.

XX AC AAC95372;

XX DT 19-FEB-2001 (first entry)

XX DE Cat flea HMT Na/H transporter cDNA ORF, SEQ ID NO:1870.

XX KW Cat flea; hindgut and Malpighian tubule nucleic acid; HMT;

XX KW flea infestation; vaccine; antiparasitic; therapeutic target;

XX KW diagnosis; detection; ss.

XX OS Ctenocephalides felis.

XX PN WO200061621-A2.

XX PD 19-OCT-2000.

XX PF 07-APR-2000; 2000WO-US09437.

XX PR 09-APR-1999; 99US-0128704.

XX PA (HESK-) HESKA CORP.

XX PI Brandt KS, Gaines PJ, Stinchcomb DT, Wisniewski N;

XX DR WPI; 2000-656323/63.

XX PT P-PSDB; AAB29621.

XX PT Flea Malpighian tubule and head and nerve cord tissue derived nucleic

XX PT infestations -

XX PS Claim 1; Page 894-895; 964pp; English.

XX CC The invention relates to novel cat flea (Ctenocephalides felis) nucleic

XX CC acids which are expressed in hindgut and Malpighian tubule (HMT) tissue

XX CC or head and nerve cord (HNC) tissue. The invention also relates to the

XX CC encoded proteins. The invention additionally encompasses expression

XX CC constructs, recombinant viruses and recombinant cells comprising the

XX CC nucleic acids of the invention, recombinant production of the proteins,

XX CC antibodies against the proteins, a method of identifying inhibitors of

XX CC the proteins, and compositions comprising the nucleic acids, and the proteins they

XX CC administration to an animal. The nucleic acids, and the proteins they

XX CC encode may be used in the prevention, treatment and diagnosis of diseases

XX CC associated with flea infestations. For example, the nucleic acids may be

XX CC used to produce an HMT or HNC protein according to standard recombinant

XX CC DNA methodology by inserting the nucleic acids into a host cell and

XX CC culturing the cell to express the protein. The HMT and HNC nucleic acids

XX CC may also be used as DNA probes in diagnostic assays (e.g., PCR) to detect

XX CC and quantitate the presence of cat flea or other homologous nucleic acid

XX CC sequences in samples. They may also be used to study the expression and

XX CC function of the proteins and their role in metabolism. The HMT and HNC

XX CC proteins may be used as antigens in the production of specific

XX CC antibodies, and in assays to identify modulators (agonists and

XX CC antagonists) of HMT and/or HNC protein expression and activity. The

XX CC anti-HMT/HNC protein antibodies and antagonists may also be used to

XX CC downregulate protein expression and activity. The antibodies may also be

XX CC used as diagnostic agents for detecting the presence of flea polypeptides

XX CC in samples (e.g., by enzyme linked immunosorbent assay (ELISA)). The

XX CC present sequence represents a cat flea HMT cDNA of the invention.

XX SQ Sequence 1824 BP; 529 A; 330 C; 355 G; 610 T; 0 other;

Query Match 6.5%; Score 141.8; DB 21; Length 1824;

Best Local Similarity 49.7%; Pred. No. 1.8e-22;

Qy 1249 gcctt 1253
||
Db 869 ATATT 865

Search completed: November 18, 2001, 11:08:22
Job time: 4566 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 18, 2001, 09:49:01 ; Search time 81 seconds

(without alignments)
5089.742 Million cell updates/sec

Title: US-09-271-584A-1

Perfect score: 2178
Sequence: 1 cctctctgttccttcctcg.....aaaaaaaaaaaaaaaaaaaaa 2178

Scoring table: IDENTITY_NUC

Gapop 10_0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued_Patents_NA:*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	53.8	2.5	2007	3	US-08-747-221B-36
2	53.8	2.5	2007	3	US-08-747-221B-38
3	53.8	2.5	2007	4	US-09-005-051-36
4	53.8	2.5	2007	4	US-09-005-051-38
5	48.8	2.2	1169	3	US-09-100-391-5
6	48.8	2.2	1302	3	US-09-100-391-11
7	48.8	2.2	2643	3	US-09-100-391-9
8	46.6	2.1	5852	1	US-07-867-106-2
9	46	2.1	3138	1	US-07-867-106-4
10	45.4	2.1	1472	4	US-08-781-420-10
11	45.4	2.1	1472	4	US-08-781-420-12
12	44.4	2.0	1368	3	US-08-874-563-5
13	44.4	2.0	1368	3	US-08-577-483-14
14	44.4	2.0	4253	3	US-08-577-483-7
15	44	2.0	6152	4	US-08-973-462-1
16	43.8	2.0	3871	2	US-08-599-455B-3
17	43.8	2.0	3871	4	US-09-069-781B-3
18	43.6	2.0	961	6	5194596-16
19	43.6	2.0	961	6	5219739-16
20	43.4	2.0	1273	1	US-08-507-431-1
21	43.4	2.0	1273	3	US-08-902-655A-1
22	43.4	2.0	1273	3	US-09-116-622-1
23	43.4	2.0	1273	4	US-09-219-277-1
24	43.4	2.0	1273	4	US-09-599-661-1
25	43.4	2.0	1393	1	US-08-174-467-18
26	43.4	2.0	1393	3	US-08-452-071-18
27	42.8	2.0	4254	2	US-08-443-639-7

28 42.6 2.0 654 5 PCT-US95-06406A-11
29 42.6 2.0 790 6 5194596-8
30 42.6 2.0 1493 6 5340934-5
31 42.6 2.0 3157 6 5198347-3
32 42.6 2.0 7218 1 US-08-232-463-14
33 42 1.9 5852 1 US-07-867-106-2
34 41.4 1.9 114 1 US-08-120-827-99
35 41.4 1.9 114 1 US-08-478-675-99
36 41.4 1.9 1279 4 US-09-277-716-31
37 41.2 1.9 2230 4 US-08-378-313-24
38 41 1.9 1939 1 US-07-715-751B-2
39 41 1.9 2133 4 US-09-187-124-1
40 40.8 1.9 1264 2 US-08-758-621-13
41 40.8 1.9 1264 4 US-09-107-858-13
42 40.8 1.9 1357 6 5340934-7
43 40.6 1.9 2444 3 US-08-906-791-1
44 40.6 1.9 11517 1 US-07-920-281C-1
45 40.6 1.9 11517 4 US-08-466-277-1

ALIGNMENTS

RESULT 1
US-08-747-221B-36
; Sequence 36, Application US/08747221B
; Patent No. 6063610
; GENERAL INFORMATION:
; APPLICANT: Silver, Gary W.
; APPLICANT: Wisniewski, Nancy
; TITLE OF INVENTION: No. 6063610el Carboxylesterase Nucleic Acid
; TITLE OF INVENTION: Molecules, Proteins and Uses Thereof
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Hesk Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Wordperfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/747,221B
; FILING DATE: No. 6063610ember 12, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: FC-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2007 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 11..1594
; US-08-747-221B-36

Query Match 2.5%; Score 53.8; DB 3; Length 2007;
Best Local Similarity 52.4%; Pred. No. 0.00029;
Matches 116; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

Qy	1954	aaattatgcttttggtaaatatccatttggtaataattgttggagacagaaaatcgt	2013
Db <th>233</th> <th>AACAAATCTGTTTAACTTGGCATTCACAGATGGTGCTACTGGCCATAATTTGT</th> <th>174</th>	233	AACAAATCTGTTTAACTTGGCATTCACAGATGGTGCTACTGGCCATAATTTGT	174
Qy <th>2014</th> <th>cctaacgcttttgagagcagaagcaaacatggcaactttgaagtgtttgatgtgatat</th> <th>2073</th>	2014	cctaacgcttttgagagcagaagcaaacatggcaactttgaagtgtttgatgtgatat	2073
Db <th>173</th> <th>CGCTCTCTTGTGAAGAACTGAACATAAAATGTGATTAAATGGAGCCGACATATTTATATTT</th> <th>114</th>	173	CGCTCTCTTGTGAAGAACTGAACATAAAATGTGATTAAATGGAGCCGACATATTTATATTT	114

Qy	1954	aaattatgccttttg	gtaaatattatccatttg	taataatgtttgtg	gagacagaatactctg	20133
Db	1775	AACAAATCTGTTT	TAGTATGCGCAT	TTCACACAGATGGT	CTACTGCGCTAAATTTGT	1834
Qy	2014	cctaacgttttg	agacagaacaaacat	ggcaacttgaagt	gtttgttatgatgat	2073
Db	1835	CGCTCTTCCTT	GAAGAACTCAACT	TAAAAATGTGATTAAT	GGAGCCACACATATTTATATTT	1894
Qy	2074	gtaattatatactat	gtgtgtttgtg	taacacaacacacacatt	gtttgttatgttttgaa	2133

Qy 2134 tttaggttttgccttcgaaaaaaaaaaaaaaaaaaaaa 2178

```

RESULT      6
US-09-100-391-11
; Sequence 11, Application US/09100391a
; Patent No. 6107088
; GENERAL INFORMATION:
; APPLICANT: Korneluk, Robert G.
; APPLICANT: Tamai, Katsuyuki
; APPLICANT: Liston, Peter
; APPLICANT: Mackenzie, Alexander E.
; TITLE OF INVENTION: XAF GENES AND POLYPEPTIDES: METHODS AND
; TITLE OF INVENTION: REAGENTS FOR MODULATING APOPTOSIS
; FILE REFERENCE: 07891/010004
; CURRENT APPLICATION NUMBER: US/09/100,391a
; CURRENT FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/052,402
; EARLIER FILING DATE: 1997-07-14
; EARLIER APPLICATION NUMBER: 60/054,491
; EARLIER FILING DATE: 1997-08-01
; EARLIER APPLICATION NUMBER: 60/056,338
; EARLIER FILING DATE: 1997-08-18
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1302
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-100-391-11

```

```

Query Match      2.2%; Score 48.8; DB 3; Length 2643;
Best Local Similarity 70.7%; Pred. No. 0.0057;
Matches 65; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 2087 tattgttttggtaacacaaactcacattgtttatgttttgaattgtgttttggct 2146
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2546 tttttattttactgttataataattattaaacttccttgtaataagaataaagtttgta 2605
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 2147 tcgaaaaaataaaaaaataaaaaaataaaaaa 2178
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2606 tggaaaaaataaaaaaataaaaaaataaaaaa 2637
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT      8
US-07-867-106-2
; Sequence 2, Application US/07867106
; Patent No. 5389526
; GENERAL INFORMATION:
; APPLICANT: Slade, Martin B
; APPLICANT: Chang, Andy C M
; APPLICANT: Williams, Keith L
; TITLE OF INVENTION: Improved Plasmid Vectors for Cellular
; TITLE OF INVENTION: Slime Moulds of the Genus Dictyostelium
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESS: Woodcock Washburn Kurtz Mackiewicz & No. 5389526ris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA

```


CITY: Fort Collins
STATE: Colorado
COUNTRY: USA
ZIP: 80525
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: WordPerfect for Windows, Version 7.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,420
FILING DATE: December 3, 1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Verser, Carol Talkington
REGISTRATION NUMBER: 37,459
REFERENCE/DOCKET NUMBER: HW-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1472 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 2..1105
PS-08-781-420-10

```

Query Match          2.1%; Score 45.4; DB 4; Length 1472;
Best Local Similarity 61.3%; Pred. No. 0.032;
Matches 73; Conservative 0; Mismatches 46; Indels 0; Gaps

QY   2060 ttgatgtatgatkgttaatttatcatcatttggtttgctgaacacaactcacattt 2111
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1341 TTAGCGCTTTTITTTGTTTACCTTATCTCTGTTATTGTACTATTTCGCCCTTG 1400

QY   2120 gtttatgtttgaattggtttgcttcgaaaaaaaaaaaaaaaaaaaaaaaaa 2178
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1401 TTTTTTAATAATTTAAATTAATTTTTTTCGAAAAAIAAAAAAAAAAAAAAAA 1459


RESULT 11
US-08-781-420-12/c
; Sequence 12, Application US/08781420
; Patent No. 6248872
; GENERAL INFORMATION:
; APPLICANT: Chandrashekar, Ramaswamy
; APPLICANT: Mehta, Kapil
; TITLE OF INVENTION: Parasitic Nematode Trangiutaminase
; TITLE OF INVENTION: Proteins, Nucleic Acid Molecules and
; TITLE OF INVENTION: Uses Thereof
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,420
; FILING DATE: December 3, 1996

```

```

/ CORRESPONDENCE ADDRESS:
/
/ ADDRESSEE: Carol Talkington Verser, Ph.D.
/ ADDRESSEE: Heskia Corporation
/ STREET: 1825 Sharp Point Drive
/ CITY: Fort Collins
/ STATE: Colorado
/ COUNTRY: USA
/ ZIP: 80525
/
/ COMPUTER READABLE FORM:
/
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: Windows 95
/ SOFTWARE: WordPerfect for Windows, Version 7.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/781,420
/ FILING DATE: December 3, 1996

```

Qy 1964 ttigtgtaattatccatttgaatatgtttgtgaggacagaaatctgtcctaagttt 2023

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 18, 2001, 08:39:10 ; Search time 1652.68 Seconds
(without alignments)
12457.533 Million cell updates/sec

Title: US-09-271-584A-1
Perfect score: 2178
Sequence: 1 cctctgtttctgttctctg.....aaaaaaaaaaaaaaaaaaaaa 2178

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues 20456230
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :	EST:*
-	1: gb_est1:*
-	2: gb_est2:*
	3: gb_est3:*
	4: gb_est4:*
	5: gb_est5:*
	6: gb_est6:*
	7: gb_est7:*
	8: gb_est8:*
	9: gb_est9:*
	10: gb_est10:*
	11: gb_est11:*
	12: gb_est12:*
	13: gb_est13:*
	14: gb_est14:*
	15: gb_est15:*
	16: gb_est16:*
	17: gb_est17:*
	18: gb_est18:*
	19: gb_est19:*
	20: gb_est20:*
	21: gb_est21:*
	22: gb_est22:*
	23: gb_est23:*
	24: gb_est24:*
	25: gb_est33:*
	26: gb_est34:*
	27: gb_est35:*
	28: gb_est36:*
	29: gb_est37:*
	30: gb_est38:*
	31: gb_est39:*
	32: gb_est40:*
	33: em_estba:*
	34: em_estfun:*
	35: em_esthum1:*
	36: em_esthum2:*
	37: em_esthum3:*
	38: em_esthum4:*
	39: em_esthum5:*
	40: em_esthum6:*
	41: em_esthum7:*
	42: em_esthum8:*
	43: em_esthum9:*
	44: em_esthum10:*
	45: em_esthum11:*
	46: em_esthum12:*
	47: em_esthum13:*
	48: em_esthum14:*
	49: em_esthum15:*
	50: em_esthum16:*
	51: em_esthum17:*
	52: em_esthum18:*
	53: em_esthum19:*
	54: em_esthum20:*
	55: em_esthum21:*
	56: em_esthum22:*
	57: em_esthum23:*
	58: em_esthum24:*
	59: em_esthum25:*
	60: em_esthum26:*
	61: em_esthum27:*
	62: em_esthum28:*
	63: em_estin1:*
	64: em_estin2:*
	65: em_estin3:*
	66: em_estin4:*
	67: em_estin5:*
	68: em_estom1:*
	69: em_estom2:*
	70: em_estov1:*
	71: em_estov2:*
	72: em_estpl1:*
	73: em_estpl2:*
	74: em_estpl3:*
	75: em_estpl4:*
	76: em_estpl5:*
	77: em_estpl6:*
	78: em_estpl7:*
	79: em_estpl8:*
	80: em_estpl9:*
	81: em_estpl10:*
	82: em_estro1:*
	83: em_estro2:*
	84: em_estro3:*
	85: em_estro4:*
	86: em_estro5:*
	87: em_estro6:*
	88: em_estro7:*
	89: em_estro8:*
	90: em_estro9:*
	91: em_estro10:*
	92: em_estro11:*
	93: em_estro12:*
	94: em_estro13:*
	95: em_estro14:*
	96: em_estro15:*
	97: em_estro16:*
	98: em_estro17:*
	99: em_estro18:*
	100: em_estro19:*
	101: em_estro20:*
	102: gb_est25:*
	103: gb_est26:*
	104: gb_est27:*
	105: gb_est28:*
	106: gb_est29:*
	107: gb_est30:*
	108: gb_est31:*
	109: gb_est32:*
	110: gb_est41:*
	111: gb_est42:*
	112: gb_est43:*
	113: gb_est44:*
	114: gb_est45:*
	115: gb_est46:*
	116: gb_est47:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est75:*
137: gb_est76:*
138: gb_est77:*
139: gb_est78:*
140: gb_est79:*
141: gb_est80:*
142: gb_est81:*
143: gb_est82:*
144: gb_est83:*
145: gb_est84:*
146: gb_est85:*
147: gb_est86:*
148: gb_est87:*
149: gb_est88:*
150: gb_est89:*
151: gb_est90:*
152: gb_est99:*
153: gb_est100:*
154: gb_est101:*
155: gb_est102:*
156: gb_est103:*
157: gb_est104:*
158: gb_est105:*
159: gb_est106:*
160: gb_est67:*
161: gb_est68:*
162: gb_est69:*
163: gb_est70:*
164: gb_est71:*
165: gb_est72:*
166: gb_est73:*
167: gb_est74:*
168: gb_est91:*
169: gb_est92:*
170: gb_est93:*
171: gb_est94:*
172: gb_est95:*
173: gb_est96:*
174: gb_est97:*
175: gb_est98:*
176: em_esthum29:*
177: em_esthum30:*
178: em_esthum31:*
179: em_esthum32:*
180: em_esthum33:*
181: em_estom3:*
182: em_estpl11:*
183: em_estro21:*
184: em_estro22:*
185: em_estro23:*
186: em_htc:*
187: gb_est107:*
188: gb_est108:*
189: gb_est109:*

190: gb_est110:*
191: gb_est111:*
192: gb_htc:*
193: em_gss_fun:*
194: em_gss_hum1:*
195: em_gss_hum2:*
196: em_gss_hum3:*
197: em_gss_hum4:*
198: em_gss_hum5:*
199: em_gss_hum6:*
200: em_gss_hum7:*
201: em_gss_hum8:*
202: em_gss_hum9:*
203: em_gss_inv1:*
204: em_gss_inv2:*
205: em_gss_inv3:*
206: em_gss_other:*
207: em_gss_pln1:*
208: em_gss_pln2:*
209: em_gss_pro:*
210: em_gss_rod1:*
211: em_gss_rod2:*
212: em_gss_rod3:*
213: em_gss_rod4:*
214: em_gss_rod5:*
215: em_gss_vrt1:*
216: em_gss_vrt2:*
217: em_gss_vrt3:*
218: gb_gss1:*
219: gb_gss2:*
220: gb_gss3:*
221: gb_gss4:*
222: gb_gss5:*
223: gb_gss6:*
224: gb_gss7:*
225: gb_gss8:*
226: gb_gss9:*
227: gb_gss10:*
228: gb_gss11:*
229: gb_gss12:*
230: gb_gss13:*
231: gb_gss14:*
232: gb_gss15:*
233: gb_gss16:*
234: gb_gss17:*
235: gb_gss18:*
236: gb_gss19:*
237: gb_gss20:*
238: gb_gss21:*
239: gb_gss22:*
240: gb_gss23:*
241: gb_gss24:*
242: gb_gss25:*
243: gb_gss26:*
244: gb_gss27:*
245: gb_gss28:*
246: gb_gss29:*
247: gb_gss30:*
248: gb_gss31:*
249: gb_gss32:*
250: gb_gss33:*
251: gb_gss34:*
252: em_gss_inv4:*
253: em_gss_rod6:*
254: em_gss_rod7:*
255: em_gss_rod8:*
256: gb_gss35:*
257: gb_gss36:*
258: gb_gss37:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SOURCE ORGANISM
barley.
Hordeum vulgare

REFERENCE AUTHORS
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae ; Triticeae; Hordeum.
1 (bases 1 to 1212)
Anderson, O.A., Appels, R., Bailey, P., Blake, T., Close, T., Cloutier, S., Dubcovsky, J., Feuillet, C., Gale, M., Graner, A., Gustafson, P., Herrmann, R.G., Holton, T., Jacquemin, J.M., Jia, J., Joudrier, P., Langridge, P., Lazo, G.R., Lin, J.J., McGuire, P., Ogihara, Y., Pecchioni, N., Qualset, C., Schuch, W., Selvaraj, G., Shariflou, M., Sorrells, M., Warburton, M. and Wenzel, G.
International Triticeae EST Cooperative (ITEC): Production of International Triticeae EST Cooperative (ITEC): Production of Expressed Sequence Tags for Species of the Triticeae
Unpublished (2000)
Contact: Herrmann RG
Botanisches Institut der LMU
Menzinger Str. 67, D-80638 Munchen GERMANY
Fax: 49 30 171683
Email: herrmann@botanik.biologie.uni-muenchen.de
International Triticeae EST Cooperative (ITEC)
http://wheat.pw.usda.gov/genome.
Location/Qualifiers
1. .1212
/organism="Hordeum vulgare"
/cultivar="Barke"
/db_xref="taxon:4513"
/clone="HW000.D12"
/tissue_type="Leaf"
/tissue_lib="ITEC HWM Barley Leaf Library"
/dev_stage="14 day old"
/note="Vector: pBluescriptSK(-); 850 bp average insert size."

BASE COUNT 296 a 294 c 276 g 346 t

ORIGIN

Query Match 19.1%; Score 415.8; DB 167; Length 1212;
-Best Local Similarity 69.1%; Pred. No. 1.5e-88;
Matches 614; Conservative 0; Mismatches 267; Indels 8; Gaps 3;

QY 587 agcagttttccgaatttcgtgactattatctgttttggctgttggaactatttt 646
DB 61 AACAAATTTCCGCAATTTTCATGACTATTACAAFTTTTCCTGGTAGTTGGGACCTTGATCT 120
QY 647 ctggcaaatcatatctctaggtgaacacagttctttaagaattggacattggaacct 706
DB 121 CCTTCAGTATATACCTTGGTGCCATGCGGCTAGTATCAAGGCTGACATAGCGCCC 180
QY 707 ttgacttggtgattattctgttattgttgccattttgtcgaacagattcagtatgta 766
DB 181 TTGAGCTTGGAGACTACCTCGCACTTGGGGCAATATTCGGCAACGGACTCTCTTCGCA 240
QY 767 cactgaggttctgaatcaagaagacagacaccttggctttacagttctgtattcgagagg 826
DB 241 CCTTGCAGGTGTTAAGCAACATGAGACACCTTTCTGTACAGTTTGGTGGTGAAG 300
QY 827 gtgtgtgaatgatgaacgtcagttgtgtcttcaacgcgattcagagcttggatctca 886
DB 301 GTGTGTTTAACGATGCGACATCAGTTGTTGTTCAATGCAATCCAGAACTTTGATCTTG 360
QY 887 ctcacctaaaccagaagctgttttcaatcttcttggaaactctgtattgtttctcc 946
DB 361 GAAATTTTCAGTAGCTCAAAATTTCTTACAAATTCATTTGGAAATTTCCCTATCTATTGGCG 420
QY 947 taagtacctgtgtgtgctgcaaccggtctgataagtcggtattgtatacaagaagctat 1006
DB 421 CCAGTACTTTCTTGGAGTATCTAGTGGACTTCTCAGTGGCTTATGTCATCAAGAACTGT 480
QY 1007 actttggaagcactcaactgacagagaggttgcccttatgtatgcttatgcttatcttt 1066
DB 481 ACTTTGGCAGGCACCTCCACTGATCGTGAAGTTGCTATTATATGATGCTCATGGCTATTAT 540

QY 1067 cttatatgttgctgagcttttgcagacttgagcgggtatctctcactgtgtttttctgtgta 1126
DB 541 CTTACATGCTGGCTGAATTCGTTGATTTGAGTGGTATTTCTACGGTTTCTTCTTGTGTA 600
QY 1127 ttgtgatgtcccaattacacatgacaaatgtaacgagagctcaagaataacaacaaagc 1186
DB 601 TTGTAATGTCGCACTATACCTGGCAATGTACAGAGAGTTCAGGGTTCACAACCAAGC 660
QY 1187 atacctttgcaactttgtcatttcttgcgagacatttattttctgtatgttggaaatg 1246
DB 661 ATGCCITGGCCATTTGTCATCTCTGAGACGTTTCTCTTCTCTATGTGTGGCATGG 720
QY 1247 atgcttggacattgacaaatgagagatccgtgagt- ----acacaccgggaacatcgat 1301
DB 721 ATGCAITTTGATATAGAGAGTGGAAATTTGGTGAACATATAGCCCAATGAATCTAT 780
QY 1302 ccagtgagctcaatcctaaagggtctgtcattggttgaagagcagcgttctgtcttcc 1361
DB 781 TGCCTTGAGCTCCATTTATTTGGCGTTGGTGTGTTGCAAGAGCTGCATTTGTTTCCC 840
QY 1362 gttatcgtttcttctaatccttagcgaagaatacaaaagcgaagaaatcaactttaacat 1421
DB 841 TCTATCTATCTCTCC--AATTGACCAAAAAAATCCAGGCGAGAGATCTCTGTTAGGCA 898
QY 1422 gcaggttgatgttggtgtgtctgtctcagagaggtgctgttatctatg 1470
DB 899 GCAAGTTATTATTGTTGGTGGGC-GGGCTCATGAGAGGGGGCGCTGTCAATG 945

RESULT 3

BE440835 502 bp mRNA EST 21-NOV-2000
LOCUS SP42f03.y1 Gm-cl043 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
DEFINITION Gm-cl043-918 5' similar to TR:004655 004655 SIMILAR TO
SODIUM/HYDROGEN EXCHANGER. ;, mRNA sequence.
ACCESSION BE440835
VERSION BE440835.1 GI:9440328
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.
1 (bases 1 to 502)
Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Corvelli, V., Khanna
A., Bolla, B., Marfa, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers
Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk
R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann
R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Genome Systems, Inc. 4633 World
Parkway Circle St. Louis, Missouri 63134 For further information
call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314)
427-3324 or contact: clones@genomesystems.com or
info@genomesystems.com web site: www.genomesystems.com
Insert Length: 1353 Std Error: 0.00
High quality sequence stop: 470.
Location/Qualifiers
1. .502
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl043-918"
/tissue_type="Hypocotyl and Plumule, germinating seeds"

FEATURES
source

QY	943	ctcctaagtactacttgcgttggtgctgcgaacgcggctctgatagaatgctgcatatgttatcaagaag	1002
Dd	301	ATCGCAAGCACATTGCTTGGGGTTTTGCACAGGTCTTTCTCAGTGCTTTACATCATTAAGAAG	360
QY	1003	ctatactcttgaaggacctcaactgaccgagaggttgcccttatgatctlatggcgctat	1062
Dd	361	CTATACATTGGCAGCACTCTACAGATCGTGAGGTGCTCTTATGATGCTAATGGCATAC	420
QY	1063	cttt 1066	
Dd	421	CTTT 424	
RESULT	6		
LOCUS	AW685820		
DEFINITION	AW685820	653 bp mRNA EST 15-JUN-2000	
ACCESSION	NF035E04NR	5' mRNA sequence.	
VERSION	AW685820		
KEYWORDS	EST.		
SOURCE	AW685820.1	GI:7560556	
ORGANISM	barrel medic.		
REFERENCE	Medicago truncatula		
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Medicago. 1 (bases 1 to 653) Watson,B.S., Shin,H.-S., Lopez-Meyer,M., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J., Flores,H.R., Inman,J.T., Weller,J.W., May, G.D. and Paiva,N.L. Expressed Sequence Tags from the Samuel Roberts Noble Foundation Medicago truncatula nodulated root library Unpublished (2000) Contact: Paiva NL Plant Biology Division The Samuel Roberts Noble Foundation 2510 Sam Noble Parkway, Ardmore, OK 73402, USA Tel: 580 221 7317 Fax: 580 221 7380 Email: nipaiva@noble.org Insert Length: 653 Std Error: 0.00 Plate: 035 row: E column: 04 Seq primer: TCACACGGAACAGCTATGAC. Location/Qualifiers 1..653 /organism="Medicago truncatula" /db_xref="taxon:3880" /clone="NF035E04NR" /clone_lib="Nodulated root" /tissue_type="root" /dev_stage="Pooled developmental" /note="Vector: Lambda zap; Four-week old Rhizobium meliloti-inoculated Medicago truncatula roots, containing a mixture of young and old roots and nodules."		
BASE COUNT	167 a	133 c	148 g
ORIGIN			205 t
Query Match	11.6%	Score 253,	DB 119;
Best Local Similarity	62.8%;	Prod. No. 8.6e-50;	Length 653;
Matches 410;	Conservative	0;	Mismatches 240;
		Indels	3; Gaps
			1;
QY	1203	gtcattcttcggagacatttcttctgttgttggaatggatgcccttgacattga	1262
Dd	1	GTCATTCAATTGCTGAGATAATTATCTTCCTTGTGTTGGATGGACTAGATATGGA	60
QY	1263	caagtggagatcgtgtaatgacacacccgggaacatcgatcgacgtgaactcaatc	1322
Dd	61	GAAGTGGCGATTTCGTPAACTCAAACCCAAAATAATCAATTGGGTCAGTTCAITGCTGTT	120
QY	1323	gggtcgttcattcgttgtaaagacagcaggttcgtcttcttcogtattcggtttctactcaactt	1382

1


```

QY 923 gaaactcttctgtattgtttcttctaagtacactctgttggtg 964
||||| ||||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 303 GAAATTTCTGTATCTATTATTATGCAAGCAACATGCTTGAG 344

RESULT 14
AW691011 654 bp mRNA EST 15-JUN-2000
LOCUS NF040D08Strf1000 Developing stem Medicago truncatula cDNA clone
DEFINITION NF040D08ST 5', mRNA sequence.
ACCESSION AW691011
VERSION AW691011.1 GI:7565747
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Medicago.
REFERENCE 1 (bases 1 to 654)
AUTHORS He,X.-Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell
,C.J., Flores,H.R., Inman,J.T., Weller,J.W., May,G.D. and Dixon
,R.A.
TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation
Medicago truncatula stem library
JOURNAL Unpublished (2000)
COMMENT Contact: Dixon RA
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7302
Fax: 580 221 7380
Email: radixon@noble.org
Insert Length: 654 Std Error: 0.00
Plate: 040 row: D column: 08
Seq primer: TCACACAGGAACAGCTATGAC.
FEATURES
source
1..654
/organism="Medicago truncatula"
/db_xref="taxon:3880"
/clone="NF040D08ST"
/clone_lib="Developing stem"
/tissue_type="stem"
/dev_stages="Pooled developmental"
/notes="Vector: Lambda Zap; Contains a mixture of
internodal stem segments"
BASE COUNT 186 a 122 c 122 g 224 t
ORIGIN

Query Match 8.9%; Score 194.6; DB 119; Length 654;
Best Local Similarity 76.9%; Pred. No. 7.1e-36;
Matches 250; Conservative 0; Mismatches 74; Indels 1; Gaps 1;

QY 294 tctctagtctgaacagctgccttgattatcgacatctgacacgtctctgttggtg 353
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 323 TCTATTGTTTCAAAACATATCAATGTTATPCCACTTCGATCATCTCTGTGTTCTAT 382

QY 354 gaactcttctgtgcaactctctgtgtgtattgttcttctgtcatcttttgggaagaa 413
||| ||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 383 GAACATGTTGTGGGACATCTCTGTGCTGTATGTCCTTGGTCATCTTCGAGGAGAA 442

QY 414 tagatggtgaacgaatccatccacgcctgtgtgattgggctagcgactggttaccat 473
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 443 TCGCTGGATGAATGAATCATCACTGCCTCTTTGATTGGTATATGCATCGTGTAATGAT 502

QY 474 ttgttgattagtaaaagaaagaaagctcgatctctctctgtctctttagtaagaatctttctt 533
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 503 TTTCGTGTTTATGTTGGTGAAGAAAGTTCACATATTCTGTTTTCAGTGAAGATCTTTTCTT 562

QY 534 catatatcttttggcaccattatattcaatg-cagggtttcaagtaaaaaaagacagt 592
||||| || ||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 563 TATATACCTGTCGCCCAATATATATTCATGCGCGGGTTTCAAGTAAAGAAAGCCATG 622

```

```

QY 593 ttctccgcaatttcgtgactattat 617
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 623 TTTTTCGCAACTGTTTCATGACTAT 647

RESULT 15
AW694621 393 bp mRNA EST 20-DEC-2000
LOCUS NF078C11Strf1085 Developing stem Medicago truncatula cDNA clone
DEFINITION NF078C11ST 5', mRNA sequence.
ACCESSION AW694621
VERSION AW694621.2 GI:11934170
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Medicago.
REFERENCE 1 (bases 1 to 393)
AUTHORS He,X.-Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell
,C.J., Flores,H.R., Inman,J.T., Weller,J.W., May,G.D. and Dixon
,R.A.
TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation
Medicago truncatula stem library
JOURNAL Unpublished (2000)
COMMENT On Apr 14, 2000 this sequence version replaced gi:7569383.
Contact: Dixon RA
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7302
Fax: 580 221 7380
Email: radixon@noble.org
Insert Length: 757 Std Error: 0.00
Plate: 078 row: C column: 11
Seq primer: TCACACAGGAACAGCTATGAC.
FEATURES
source
1..393
/organism="Medicago truncatula"
/db_xref="taxon:3880"
/clone="NF078C11ST"
/clone_lib="Developing stem"
/tissue_type="stem"
/dev_stages="Pooled developmental"
/notes="Vector: Lambda Zap; Contains a mixture of
internodal stem segments"
BASE COUNT 111 a 94 c 72 g 116 t
ORIGIN

Query Match 8.8%; Score 192.6; DB 119; Length 393;
Best Local Similarity 71.0%; Pred. No. 2e-35;
Matches 255; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 1288 ccgggaacatcgatcgagtcagtcacttaacttaactggttcgttcgcatggttggaagagca 1347
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 5 CCTGGAACATCTATAGCAGCAAGTTCAGTATTGTTGGTCTATACTCTTCTGGAAGAGCA 64

QY 1348 gcttcgctcttcctcgcttatctggtttctatcttaacttagccaagaagaatcaaaagcgagaa 1407
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65 GCGTTTGTTCCTTCCTTATCTCTTATCCACTTGACTTAAAAATCACAACATCAGAAG 124

QY 1408 atcaactttaacatcgaggttgattggttcgttcgttcctcatgagaggtctgtatct 1467
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 125 ATCTCCTTCAGCAGCAAGTGATCATTTGGTGGCGTGGCTTATGAGAGGTGCTGTTTCA 184

QY 1468 atggctcttgcatacaacaagtttaacaaggccggcgacacacagatgtacgcgggaatgca 1527
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 185 ATGGCACTTTCGTTATATCACTAGTTCACCATGTCGGGGCATACTCACTGCGTAGCAATGCA 244

QY 1528 atcatgatcacgagtagcagataaactgtctcttcttttagcacagtggtgttggatgtg 1587
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 245 ATCATGATACCAAGCAGCCACTGTTGCTCTTTTCAGCAGCAGTGGTGTGTTGGTTGCTG 304

```

Qy 1588 accaaacacataagctacacctattaccgcaccagacgcccacacgagcaigtgtatc 1646
Db 305 ACTAAGCCACTTATAAGGCTTCTACTACCTCATCCTTAAATCACAAGCAGCATGACAAC 363

Search completed: November 18, 2001, 10:14:36
Job time: 5726 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2001, 13:26:37 ; Search time 15.28 Seconds
(without alignments)
792.328 Million cell updates/sec

Title: US-09-271-584A-2

Perfect score: 2755
Sequence: 1 MLDSLVKLSLSTSDHASV.....FVFPVGSPTERNPPDLKA 538

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA: *
1: /cgn2.6/ptodata/2/iaa/5A_COMB.pep: *
2: /cgn2.6/ptodata/2/iaa/5B_COMB.pep: *
3: /cgn2.6/ptodata/2/iaa/6A_COMB.pep: *
4: /cgn2.6/ptodata/2/iaa/6B_COMB.pep: *
5: /cgn2.6/ptodata/2/iaa/PTUS_COMB.pep: *
6: /cgn2.6/ptodata/2/iaa/backfiles1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	471	17.1	832	2	US-08-677-734A-12
2	467	17.0	831	2	US-08-677-734A-11
3	462	16.8	834	2	US-08-677-734A-9
4	462	16.8	834	2	US-08-677-734A-10
5	104	3.8	429	2	US-08-677-049-5
6	99	3.6	462	3	US-09-238-796-2
7	98.5	3.6	397	1	US-08-098-141-2
8	95	3.4	508	1	US-08-472-028A-4
9	95	3.4	508	2	US-08-808-931-4
10	95	3.4	508	3	US-08-808-323-4
11	95	3.4	508	3	US-09-050-603A-4
12	95	3.4	508	3	US-09-102-420B-4
13	95	3.4	508	4	US-09-071-296-4
14	95	3.4	508	4	US-09-196-268-4
15	95	3.4	508	4	US-09-015-683-4
16	95	3.4	508	4	US-09-191-998-4
17	93	3.4	342	3	US-08-852-824-2
18	93	3.4	694	2	US-08-895-522-4
19	93	3.4	694	3	US-09-193-391-4
20	92.5	3.4	649	4	US-08-800-291B-5
21	92.5	3.4	650	4	US-08-800-291B-4
22	92.5	3.4	1711	3	US-08-369-822C-10
23	92.5	3.4	1711	3	US-08-582-776C-10
24	92.5	3.4	1711	3	US-08-434-831B-10
25	91.5	3.3	325	1	US-08-118-270-51
26	91.5	3.3	325	5	PCR-US93-08528-51
27	91.5	3.3	649	4	US-08-800-291B-6

28 90.5 3.3 635 2 US-09-014-969-11
29 90 3.3 359 1 US-08-148-209A-4
30 89.5 3.2 1059 4 US-09-134-513-2
31 88.5 3.2 341 2 US-08-846-762-92
32 88.5 3.2 471 1 US-07-996-772A-11
33 88.5 3.2 471 4 US-09-032-742-2
34 88.5 3.2 616 1 US-08-149-100-2
35 88 3.2 459 4 US-09-097-889-22
36 88 3.2 549 2 US-08-676-279-59
37 87.5 3.2 471 4 US-09-032-742-11
38 87.5 3.2 471 4 US-09-032-742-14
39 87.5 3.2 707 2 US-08-576-165-4
40 87.5 3.2 764 1 US-08-424-567-2
41 87.5 3.2 764 2 US-08-711-928-2
42 87.5 3.2 764 4 US-09-184-937-2
43 87 3.2 576 2 US-08-676-279-58
44 86.5 3.1 471 4 US-09-032-742-17
45 86 3.1 672 1 US-07-841-651-2

ALIGNMENTS

RESULT 1
US-08-677-734A-12
; Sequence 12, Application US/08677734A
; Patent No. 5871919

; GENERAL INFORMATION:

; APPLICANT: Brant, Steven R.

; APPLICANT: Yun, Chris C.H.

; APPLICANT: Donowitz, Mark

; APPLICANT: Tse, Chung-Ming

; TITLE OF INVENTION: Cloning, Tissue Distribution, and

; TITLE OF INVENTION: Functional Analysis Of The Human Na+/H+ Exchanger Isoform,

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &

; STREET: 1300 I Street, N.W., Suite 700

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20005-3315

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/677,734A

; FILING DATE: 10-JUL-1996

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Fordis, Jean B.

; REGISTRATION NUMBER: 32,984

; REFERENCE/DOCKET NUMBER: 05387.0043-00000

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 408-4000

; TELEFAX: (202) 408-4400

; INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 832 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; US-08-677-734A-12

Query Match 17.1%; Score 471; DB 2; Length 832;
Best Local Similarity 29.1%; Pred. No. 4.8e-39;
Matches 148; Conservative 96; Mismatches 197; Indels 68; Gaps 19;


```

; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/677,734A
; FILING DATE: 10-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fordis, Jean B.
; REGISTRATION NUMBER: 32,984
; REFERENCE/DOCKET NUMBER: 05387.0043-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 834 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-677-734A-9

```

```

Query Match 16.8%; Score 462; DB 2; Length 834;
Best Local Similarity 28.6%; Pred. No. 4e-38;
Matches 146; Conservative 97; Mismatches 203; Indels 64; Gaps 18;

```

```

Qy 20 VVALNLFVALLCACIVLGHLEENRMNESITALLIGLGTGVTILLISKG---KSSHLV 76
Db 57 VIALWILVASLAK---IGFHLSHKVTSVVPESALLIVLG-----LVLGGIVWAADHTAS 107
Qy 77 F--SEDLFFIYLLPPIIFNAGFOVKKQFFRNFVIMLFGAVGTIISCTIISLGVTOFFK 134
Db 108 FTLTPVTFVFFLLPPIIVLDAGYFMPNRLFGNLGTLILYAVVGTVMNAATTGLSLYGVFL 167
Qy 135 KLDIGTFDLG--DYLAIGAIFAATDSVCTLOVLNQ--DETPLLISLVFGGVNDATSVVV 191
Db 168 SGLMGDLQIGLDFLLFGSLMAADVPVAVLAVFEVHVNEVLFIIVFGESLNDATVVL 227
Qy 192 FNAIQSFDTLHLNHEAFLHLLGNFLYLLSTLLGAATGLISAYVKKL-YFGRHSTDR 250
Db 228 YNPFESVALGGDNVTGDCVKGIIVFFVVS-LGDTLVGVVFAFLSLVTRFKHVRIE 286
Qy 251 VALMMLMAYLSYMLAEFLDGLTFFCGIVMNSHYTHNVTESRITTKKTEATLSFLA 310
Db 287 PGFVFIISYLSYLTSEMLSALAITFCGICCKYKVNANISEQATVRYTMKMLASSA 346
Qy 311 ETFLFYVGMADLIDKWRVSPTGTSIAVSSILMGLVMGAAAFVPLSFLSLAKKN 370
Db 347 ETIIFMFLGISAVNPFIW-----TWNTAFVLLTV--FISVYRAIGVVLQTLWLLNRYRMV 399
Qy 371 QSEKINFNMQVVIWWSGLMRGAVSMALAYNKFTKTRAGTDDVRGNATIMTSTITVCLFSTVV 430
Db 400 QLEPID---QVLSYGGGL-RGAVAFALV---VLLDGDVKVEKNLFVSTTIIVFFTVIF 451
Qy 431 FGMUTKPLISYLLPHONATTSMISDNTPK---SIHPLLDQDSPTIEPSGNHNVRPDSI 487
Db 452 QGLTIKPLVQWLKVR-----SEHREPLNEKLHGRAFD-----HILSAIEDI 494
Qy 488 RGLFTRPTRVHYV---WROFDDSMRPV 513
Db 495 SG-----QIGHNVLKDKWSHFDRKFLSRV 518

```

RESULT 4

```

; US-08-677-734A-10
; Sequence 10, Application US/08677734A
; Patent No. 5871919
; GENERAL INFORMATION:

```

```

; APPLICANT: Brant, Steven R.
; APPLICANT: Yun, Chris C.H.
; APPLICANT: Donowitz, Mark
; APPLICANT: Tse, Chung-Ming
; TITLE OF INVENTION: Cloning, Tissue Distribution, and
; TITLE OF INVENTION: Functional Analysis Of The Human Na+/H+ Exchanger Isoform,
; TITLE OF INVENTION: NHE3.
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/677,734A
; FILING DATE: 10-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fordis, Jean B.
; REGISTRATION NUMBER: 32,984
; REFERENCE/DOCKET NUMBER: 05387.0043-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 834 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-677-734A-10

```

```

Query Match 16.8%; Score 462; DB 2; Length 834;
Best Local Similarity 28.6%; Pred. No. 4e-38;
Matches 146; Conservative 97; Mismatches 203; Indels 64; Gaps 18;

```

```

Qy 20 VVALNLFVALLCACIVLGHLEENRMNESITALLIGLGTGVTILLISKG---KSSHLV 76
Db 57 VIALWILVASLAK---IGFHLSHKVTSVVPESALLIVLG-----LVLGGIVWAADHTAS 107
Qy 77 F--SEDLFFIYLLPPIIFNAGFOVKKQFFRNFVIMLFGAVGTIISCTIISLGVTOFFK 134
Db 108 FTLTPVTFVFFLLPPIIVLDAGYFMPNRLFGNLGTLILYAVVGTVMNAATTGLSLYGVFL 167
Qy 135 KLDIGTFDLG--DYLAIGAIFAATDSVCTLOVLNQ--DETPLLISLVFGGVNDATSVVV 191
Db 168 SGLMGDLQIGLDFLLFGSLMAADVPVAVLAVFEVHVNEVLFIIVFGESLNDATVVL 227
Qy 192 FNAIQSFDTLHLNHEAFLHLLGNFLYLLSTLLGAATGLISAYVKKL-YFGRHSTDR 250
Db 228 YNPFESVALGGDNVTGDCVKGIIVFFVVS-LGDTLVGVVFAFLSLVTRFKHVRIE 286
Qy 251 VALMMLMAYLSYMLAEFLDGLTFFCGIVMNSHYTHNVTESRITTKKTEATLSFLA 310
Db 287 PGFVFIISYLSYLTSEMLSALAITFCGICCKYKVNANISEQATVRYTMKMLASSA 346
Qy 311 ETFLFYVGMADLIDKWRVSPTGTSIAVSSILMGLVMGAAAFVPLSFLSLAKKN 370
Db 347 ETIIFMFLGISAVNPFIW-----TWNTAFVLLTV--FISVYRAIGVVLQTLWLLNRYRMV 399
Qy 371 QSEKINFNMQVVIWWSGLMRGAVSMALAYNKFTKTRAGTDDVRGNATIMTSTITVCLFSTVV 430
Db 400 QLEPID---QVLSYGGGL-RGAVAFALV---VLLDGDVKVEKNLFVSTTIIVFFTVIF 451

```

Query Match 3.8%; Score 104; DB 2; Length 429;
Best Local Similarity 21.5%; Pred. No. 0.034;
Matches 105; Conservative 67; Mismatches 189; Indels 128; Gaps

```

60  ICKKIPAYLGSS-----FATISPVULLLPJGYEV-----ALGCFIMCGV 99
125  ISLGWTOFFKKLDIGTFD-LGDYLAIGAIFAATDSVCTQLVINODETPLLYSLVFGEGVV 183
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
100  LFCIVSFIVKKAGCTGWLDPVPPAAMGAIVA---VIGLELAG---VAAGMAGLLPAEGQT 153
194  NDATSVVVFNATOSFDLPHLHNEARAFHLLGNFL---YLFLLSTILCAATGLISAYVKKL 240
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
154  PDSKTHII-----SITFL---AVTLGVSLFRGLAIIPIILGVLGVY-----AL 195
241  YFGRHSSTDREVALMLMAYLSYMLAEFLDSLITVFFCG-IWMSHYTHWNVTSSRITT 299
196  SFAMGIVDTPIINAHWFALPTLYTPREWFALITLPAALVVAIEAHVGHVLTVANIVKK 255
300  -----KHTFAT-LSPLAETFTIFYVGMDALDIDKWSRVSDDPGTS-----IAVS 342
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
256  DLLRDPGLHRSMEFANGSLTSTVSGFF-----GSTPNTTYGENIGVMAIT 298
343  SILMGLVWVCRAAFVPPISFLSNLAKKNQ-----SEKINFNQVVIVWSGLMKGAVSMA 396
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
299  RYVSTWVIGGAIAIFALLSCVCKLAAAIQMIPLVPMGGVSLLYGVIGASGI-RVLTESK 357
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
397  LAYNKTRAGTHTDVRCNAIMTSTITV-----CLFSTVVFQM-LTKP 437
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
358  VDYNK-----AQNILTSVLLIIGVSGAKVIGRAELKGMALATIVIGLSLIFK 407
438  LISYLLPHQ 446
   :  :  :  :  :  :
408  LISVLRPEE 416
   :  :  :  :  :  :

```

3300 -----KHTFAT--LSPFLAETIFLYVGMDALIDKWRVSVDTPGTS-----IAVS 3424
3301 SFAGSIVDIIPIILINGNWAFLPILIPREFEALILDFRALVYIGRVGRLVIANIVAK 2302
2556 DLLRDPGLHRSMFANGLSIVISGFF-----GSTPNTTYGENIGVVAIT 2988
343 SILMGLVWVGRAAFVPEPLSFLNLAKNQ-----SEKINFNQVVIVWMSGLMRGAYSMA 3966
299 RVYSTWVIGGAIAFALLSCYCKLAAAIQMTPLPVWGGVSYLLLYGVICASGI--RVLIIESK 357
337 LAYNKTFRAGTHTDVNRNATIMTSTIVT-----CLFSTVVVFCM--LTKP 437
358 VDYNK-----AQNLILTSVILIIGYSGAKVNIAGAEELKGMALATIVIGLSLIFK 407
438 LISYLLPHQ 446
408 LISVLRPEE 416

```

      :          |          |          |          |          |          |
299 RVSTWIGGAIAFALLSCVGLKAAAIQIPLVMGGVSLLLYGVIGASGI-RVLIESK 357
      :          |          |          |          |          |          |
397 LAYNKFTRAGTVDVRGNAMITMTSTTV-----CLFSTVFVGM-LTKP 437
      :          |          |          |          |          |          |
358 VDYNNK-----AQNLILTSVIIIGSGAKVNIGAEELKGMALETIVGIGLSLIFK 407
      :          |          |          |          |          |          |
438 LISYLLPHQ 446
      :          |          |          |          |          |          |
408 LISVLRPEE 416

```

3-09-238-796.-2
Sequence 2, Application US/09238796
Patent No. 6074845
GENERAL INFORMATION:
APPLICANT: AIYAR, NAMBI V.
APPLICANT: DISA, JYOTI
TITLE OF INVENTION: BECLR:BOVINE CALCITONIN RECEPTOR-LIKE
TITLE OF INVENTION: RECEPTOR
NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:
ADDRESSEE: Ratner & Prestia
STREET: P.O. Box 980
CITY: Valley Forge
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/238,796
FILING DATE: 28-JAN-1999
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:

NAME: Prestia, Paul F
REGISTRATION NUMBER: 23,031

```

; REFERENCE/DOCKET NUMBER: GP-70599
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0700
; TELEFAX: 846169
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 462 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-238-796-2

```

```

Query Match          3.6%; Score 99; DB 3; Length 462;
Best Local Similarity 20.9%; Pred. No. 0.12;
Matches 90; Conservative 59; Mismatches 130; Indels 152; Gaps 23;

```

```

QY 13 STSDHASY-VALNLFVALLCAGIVLGHLEENRWNESITALLIGLGTGVITILLISGKS 71
DB 129 NVNTHKEVKVTKALNFY-----LTIIGHVL-----SIASLLISLG-----IFFYFKSL 171
QY 72 SHLVFSEDLFFIYLLPPIIFNAGFQVKKOFFERFVITMLFGAVG--TIIISCTIISLG 128
DB 172 QORITLHKLNLFVFCNSVI-----TIIHLTAVANNOALVATNPVSCK 214
QY 129 VTOFFKKLDIGTDLGDLAIGAIFAATDSVCTQLVQNODETPLLISLVFGEGV-VNDAT 187
DB 215 VSQFIHLXLMG-----CN-----YFWMCEGIVLHFLV 242
QY 188 SVVFNRIQSFDLTHLNEAFAHLLGNFLYFLLSLILCAATGLISA---YVKKLYEGR 244
DB 243 VVAVFAEKQ-----HLM---WYIF-----LGMGFPLIPACIHAVARRLYND 281
QY 245 H---STDREVALMMLMAYLSYLAELFDLSGILTVFFCGIVMSHYTHNVTESSRIITKH 301
DB 282 NCWISSDTQLLYIHGPICALLVNLFFLLNIVRVLITKLKVTQHAESNLYMKA---VRA 338
QY 302 TFATLSFLAETFLVGMGDAIDDKRSVSDTPTGTSIA-----VSSILM---GLVMVG 352
DB 339 TLILVPLLGIEFVLI-----PWR-----PEGKAEIYDIYIIMLMHYOGLLV-- 381
QY 353 RAAVFPPLSPLSLNAKKNOSE-KINFNMQVVIWISGLARGAVSMALAYNKETFRAGHTDVR 411
DB 382 STIFCFEFGVEQAILRRNNQYKIQFG-----NNFS---HSDTL 417
QY 412 GNAMITSTIT 422
DB 418 RSASYTVVSTIS 428

```

```

RESULT 7
US-08-098-141-2
; Sequence 2, Application US/08098141
; Patent No. 5441875
; GENERAL INFORMATION:
; APPLICANT: Hediger Ph.D., Matthias A.
; TITLE OF INVENTION: Urea Transporter Polypeptide
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Choate, Hall & Stewart
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02119
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/098.141
; FILING DATE: 19930723
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kaplan Esq., Warren A.
; REGISTRATION NUMBER: 34,199
; REFERENCE/DOCKET NUMBER: 092662-001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-227-5020
; TELEFAX: 617-227-7566
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 397 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-098-141-2

```

```

Query Match          3.6%; Score 98.5; DB 1; Length 397;
Best Local Similarity 21.3%; Pred. No. 0.11;
Matches 68; Conservative 49; Mismatches 73; Indels 129; Gaps 17;

```

```

QY 34 IVLGHLEENRWNESITALLIGLGTGVITILLISGKS-----H-----LLVFS 78
DB 80 IVIG-LFVNQPMW--AIACLGTVMTSTLALLISQDRASIASGLHGYNGVLVGLLIAVFS 136
QY 79 E--DLEFFIYLLPPIIFNAGFQVKKOFFERFVITMLFGAVGTIISCTIISLGVTQFFKKL 136
DB 137 DKGDYVWLLLPVIMS-----MSCPLSSALGTIFS 170
QY 137 DIGTFDLGDLAIGAIFAATDSVCTQLVQNODETPLLISLVFGEGVVDATSV--VVFNA 194
DB 171 DUPVFTLPENIATVLYLAATGH-----YNLFFPTTLQPVSSVPNITWSE 215
QY 195 IQSFDLTHLNEAFAHLL-----GNFLY-LFLLSTL-----LGA 227
DB 216 IQ-----VPLLRLRAIPVGIGVYGCNDPWTGGIFLIALFISSPLICLHAAIGS 263
QY 228 ATCLISAYVI---KKLYEGRHSTDREVA-----LMLMAYLSYMLA---ELF----- 268
DB 264 TGMGLAALTIATPDSIYFGLCGFNSTLACIAGVGFMFTVITWQTHLLAVACALFAAYVGA 323
QY 269 DLSGILTVF-----FC 279
DB 324 ALTNVLSVFLPTCTWPEC 342

```

```

RESULT 8
US-08-472-028A-4
; Sequence 4, Application US/08472028A
; Patent No. 5767373
; GENERAL INFORMATION:
; APPLICANT: Ward, Eric R
; APPLICANT: Volrath, Sandra
; TITLE OF INVENTION: Manipulation of Protoporphyrinogen
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472.028A
; FILING DATE:

```

```

; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: CGC 1748/CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8614
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 508 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-472-028A-4

Query Match          3.4%; Score 95; DB 1; Length 508;
Best Local Similarity 18.5%; Pred. No. 0.36;
Matches 55; Conservative 48; Mismatches 87; Indels 108; Gaps 15;

QY 290 NVTESRIITKHTFATLSFLAETFIIFYVG----MDALDIDKRSVSDTPGTSIAVSSIL 345
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 SAADPDSLMSKHSFPDLNVEKSGSIIVGAIKTKFAAKGKSRDTKSSPGTKKG----- 235
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 346 MGLVMVGRAAFVFP--LSFL-SNLAKKNQSEKINFNMQVVIWWSGLMRGAVSMALAYNKF 402
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 236 -----SRGSFSGKGMQILPDTLCKLSLHDEINLDSKV-----LSLSYNSG 276
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 403 TR-----AGHTDVRG-----NAIMITSTITVC-----424
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 277 SRQENWSLSCVSHNETORQNPYDAVIMTAPL--CNVKEMKVMKGGQPPQLNPLPEINYM 334
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 425 ---LFTSVVFGMLTKPLISY--LLPHQNTAT-----TSMSSDNDTPKSIHIPLLD 468
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 335 PLSVLITTTFKKVKRPLEGFGVLIPSKQKHGFKTLGTLFSSMMFPDRSPSDVHL----- 390
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 469 QDSFIEPSGNHNVP--PDSIRGELT-----RPTRTVHYWRO-----FDDSF 509
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 391 YTTFIGGSRNOELAKASTDELKQVTSDLQRLLGVEGEPVSVNHYWYRKAFPLYDSSY 448
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 9
US-08-808-931-4
; Sequence 4, Application US/08808931
; Patent No. 5939602
; GENERAL INFORMATION:
; APPLICANT: Johnson, Marie
; APPLICANT: Volrath, Sandra
; APPLICANT: Potter, Sharon
; APPLICANT: Ward, Eric
; APPLICANT: Helfetz, Peter
; TITLE OF INVENTION: DNA Molecules Encoding Plant
; TITLE OF INVENTION: Protoporphyryinogen Oxidase and Inhibitor-Resistant Mutants
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5939602artis Corporation
; STREET: 520 White Plains Road, P.O. Box 2005
; CITY: Tarrytown
; STATE: NY
; COUNTRY: USA
; ZIP: 10591-9005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,931
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: US 60/012,705
; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/013,612
; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/020,003
; FILING DATE: 21-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Weigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC 1847
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8587
; TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 508 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-808-931-4

Query Match          3.4%; Score 95; DB 2; Length 508;
Best Local Similarity 18.5%; Pred. No. 0.36;
Matches 55; Conservative 48; Mismatches 87; Indels 108; Gaps 15;

QY 290 NVTESRIITKHTFATLSFLAETFIIFYVG----MDALDIDKRSVSDTPGTSIAVSSIL 345
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 SAADPDSLMSKHSFPDLNVEKSGSIIVGAIKTKFAAKGKSRDTKSSPGTKKG----- 235
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 346 MGLVMVGRAAFVFP--LSFL-SNLAKKNQSEKINFNMQVVIWWSGLMRGAVSMALAYNKF 402
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 236 -----SRGSFSGKGMQILPDTLCKLSLHDEINLDSKV-----LSLSYNSG 276
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 403 TR-----AGHTDVRG-----NAIMITSTITVC-----424
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 277 SRQENWSLSCVSHNETORQNPYDAVIMTAPL--CNVKEMKVMKGGQPPQLNPLPEINYM 334
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 425 ---LFTSVVFGMLTKPLISY--LLPHQNTAT-----TSMSSDNDTPKSIHIPLLD 468
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 335 PLSVLITTTFKKVKRPLEGFGVLIPSKQKHGFKTLGTLFSSMMFPDRSPSDVHL----- 390
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 469 QDSFIEPSGNHNVP--PDSIRGELT-----RPTRTVHYWRO-----FDDSF 509
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 391 YTTFIGGSRNOELAKASTDELKQVTSDLQRLLGVEGEPVSVNHYWYRKAFPLYDSSY 448
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
US-08-808-323-4
; Sequence 4, Application US/08808323
; Patent No. 6018105
; GENERAL INFORMATION:
; APPLICANT: Johnson, Marie
; APPLICANT: Volrath, Sandra
; APPLICANT: Ward, Eric
; TITLE OF INVENTION: Promoters from Plant
; TITLE OF INVENTION: Protoporphyryinogen Oxidase Genes
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6018105artis Corporation
; STREET: 520 White Plains Road, P.O. Box 2005
; CITY: Tarrytown
; STATE: NY
; COUNTRY: USA
; ZIP: 10591-9005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,931
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

```


APPLICATION NUMBER: US 08/808,323
FILING DATE: 28-FEB-1996
CLASSIFICATION: 800
PRIOR APPLICATION NUMBER: US 60/012,705
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,612
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/020,003
FILING DATE: 21-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC 1846
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 508 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-808-323-4

Query Match 3.4%; Score 95; DB 3; Length 508;
Best Local Similarity 18.5%; Pred. No. 0.36;
Matches 55; Conservative 48; Mismatches 87; Indels 108; Gaps 15;

QY 290 NVTSSRTTKHTPATLSFLAETFIPLYG----MDALDIDKWSVSDTPGTSLVSSIL 345
DB 181 SAADPDSLMSKHSFPDLWNVEKSGSIIVGAIRTKFAAGKGRDRTKSSPGTKG----- 235
QY 346 MGLVMVGRAAFVFP--LSFL-SNLAKKNOSEKINFNQVVMWGLMRGAVSMALAYNKF 402
DB 236 -----SRGSEFSGKGMQILPDLTKLSLHDEINLDSKV-----LSLSYNSG 276
QY 403 TR-----AGHTDVRG-----NAIMTSTITVC----- 424
DB 277 SRQENWLSVCVSHNETQRQPHYDAVIMTAPL--CNVKEMKVMKGQPFQNLNPLPEINYM 334
QY 425 ----LFSTVVFGLTKPLISY--LLPHONAT-----TSMLSDDNTPKSIHIPLLD 468
DB 335 PLSVLITFTTEKVKRPLEGFGVLPSKEQKHGFKTLGTLFSSMMFPDRSPSDVHL---- 390
QY 469 QDSFIEPSGNHNVP--PDSIRGFLT-----RPTRTVHYWRO-----FDDSF 509
DB 391 YTTFIGGSRNOELAKASTDELKQVWTSDLQRLGLGVEGEPVSVNHYWYRKAFPLYDSSY 448

RESULT 11
US-09-050-603A-4
Sequence 4, Application US/09050603A
Patent No. 6023012
GENERAL INFORMATION:
APPLICANT: Volrath, Sandra
APPLICANT: Johnson, Marie
APPLICANT: Potter, Sharon
APPLICANT: Ward, Eric
APPLICANT: Heifetz, Peter
TITLE OF INVENTION: DNA Molecules Encoding Plant
TITLE OF INVENTION: Protoporphyrinogen Oxidase
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESS: NO. 6023012artis Corporation
STREET: 3054 Cornwalis Road
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/050,603A
FILING DATE: 30-MAR-1998
CLASSIFICATION: 800

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/808,931
FILING DATE: 28-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/012,705
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,612
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/020,003
FILING DATE: 21-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC 1847
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 508 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-050-603A-4

Query Match 3.4%; Score 95; DB 3; Length 508;
Best Local Similarity 18.5%; Pred. No. 0.36;
Matches 55; Conservative 48; Mismatches 87; Indels 108; Gaps 15;

QY 290 NVTSSRTTKHTPATLSFLAETFIPLYG----MDALDIDKWSVSDTPGTSLVSSIL 345
DB 181 SAADPDSLMSKHSFPDLWNVEKSGSIIVGAIRTKFAAGKGRDRTKSSPGTKG----- 235
QY 346 MGLVMVGRAAFVFP--LSFL-SNLAKKNOSEKINFNQVVMWGLMRGAVSMALAYNKF 402
DB 236 -----SRGSEFSGKGMQILPDLTKLSLHDEINLDSKV-----LSLSYNSG 276
QY 403 TR-----AGHTDVRG-----NAIMTSTITVC----- 424
DB 277 SRQENWLSVCVSHNETQRQPHYDAVIMTAPL--CNVKEMKVMKGQPFQNLNPLPEINYM 334
QY 425 ----LFSTVVFGLTKPLISY--LLPHONAT-----TSMLSDDNTPKSIHIPLLD 468
DB 335 PLSVLITFTTEKVKRPLEGFGVLPSKEQKHGFKTLGTLFSSMMFPDRSPSDVHL---- 390
QY 469 QDSFIEPSGNHNVP--PDSIRGFLT-----RPTRTVHYWRO-----FDDSF 509
DB 391 YTTFIGGSRNOELAKASTDELKQVWTSDLQRLGLGVEGEPVSVNHYWYRKAFPLYDSSY 448

RESULT 12
US-09-102-4208-4
Sequence 4, Application US/091024208
Patent No. 6084155
GENERAL INFORMATION:
APPLICANT: Volrath, Sandra
APPLICANT: Johnson, Marie
APPLICANT: Ward, Eric
APPLICANT: Heifetz, Peter
TITLE OF INVENTION: HERBICIDE-TOLERANT PROTOPORPHYRINOGEN
TITLE OF INVENTION: OXIDASE ("PROTOX")


```

QY 469 QDSFTFPGNNHNPV--PDSIRGFLT-----RPTRTVHYWRQ---FDDSF 509
      :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
DB 391 VTTTIGGRNOELAKASTDELQVVTSLDRLLCVGECPVSNHYWTKAPLPDSSY 448

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Search completed: November 17, 2001, 13:27:59
Job time: 82 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2001, 13:26:57 ; Search time 27.58 Seconds
(without alignments)
2580.859 Million cell updates/sec

Title: US-09-271-584A-2
Perfect score: 2755
Sequence: 1 MDSLVSKLPSLSTSDHASV.....FVPFVPGSPTEPNPPDLKA 538

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_16:*

1: sp.archaea:*

2: sp.bacteria:*

3: sp.fungi:*

4: sp.human:*

5: sp.invertebrate:*

6: sp.mammal:*

7: sp.mhc:*

8: sp.organelle:*

9: sp.phage:*

10: sp.plant:*

11: sp.rodent:*

12: sp.unclassified:*

13: sp.vertebrate:*

14: sp.virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2755	100.0	538	10	Q9ZPK3	Q9ZPK3 arabidopsis
2	2168.5	78.7	555	10	Q9F2N0	Q9F2N0 atriplex gm
3	2136	77.5	457	10	Q04655	Q04655 arabidopsis
4	2041	74.1	542	10	Q9F8R3	Q9F8R3 pharbitis n
5	2023.5	73.4	535	10	Q9SXJ8	Q9SXJ8 oryza sativ
6	1910.5	69.3	552	10	Q9SQ00	Q9SQ00 arabidopsis
7	1567	56.9	529	10	Q9FJ63	Q9FJ63 arabidopsis
8	615.5	22.3	561	5	Q9XZM4	Q9XZM4 drosophila
9	615.5	22.3	649	5	Q9VPEJ1	Q9VPEJ1 drosophila
10	586	21.3	140	10	Q9LKH5	Q9LKH5 mesembryant
11	583.5	21.2	703	5	Q9XW14	Q9XW14 caenorhabdi
12	576.5	20.9	616	4	Q75827	Q75827 homo sapien
13	570	20.7	727	5	Q9WM99	Q9WM99 drosophila
14	568	20.6	687	5	Q90624	Q90624 drosophila
15	544	19.7	569	3	Q13726	Q13726 schizosacch
16	542	19.7	153	10	Q9LKH6	Q9LKH6 mesembryant
17	531	19.3	629	5	Q20944	Q20944 caenorhabdi
18	494.5	17.9	437	4	Q9V507	Q9V507 homo sapien
19	489.5	17.0	560	3	Q9HEX3	Q9HEX3 pneumocysti

20	458.5	16.6	518	5	Q9VIF9	Q9VIF9 drosophila
21	451	16.4	411	4	Q9Y2E8	Q9Y2E8 homo sapien
22	451	16.4	1203	5	Q9NGZ4	Q9NGZ4 drosophila
23	442.5	16.1	1179	5	Q9NCQ0	Q9NCQ0 aedes aegypt
24	432	15.7	673	5	Q23706	Q23706 carcinus ma
25	424	15.4	478	5	Q01995	Q01995 caenorhabdi
26	414	15.0	698	13	Q9W724	Q9W724 cyprinus ca
27	414	15.0	813	13	Q9W6Q1	Q9W6Q1 amphiuma tr
28	411.5	14.9	779	13	Q9W714	Q9W714 platichthys
29	410	14.9	781	13	P70009	P70009 xenopus lae
30	406	14.7	798	5	Q21386	Q21386 caenorhabdi
31	391.5	14.2	634	5	Q9T2H0	Q9T2H0 caenorhabdi
32	386.5	14.0	602	5	Q09432	Q09432 caenorhabdi
33	385.5	14.0	651	5	Q16452	Q16452 caenorhabdi
34	366.5	13.3	660	5	Q23617	Q23617 caenorhabdi
35	315.5	11.5	768	13	Q9W7S0	Q9W7S0 anguilla an
36	311.5	11.3	494	1	Q29412	Q29412 archaeoglob
37	289	10.5	684	5	Q19444	Q19444 caenorhabdi
38	279.5	10.1	531	10	Q9WA14	Q9WA14 arabidopsis
39	278.5	10.1	527	2	P73863	P73863 synechocyst
40	269	9.8	524	2	Q9LCB5	Q9LCB5 bacillus su
41	259	9.4	530	2	Q9F3L8	Q9F3L8 streptomyce
42	258	9.4	1146	10	Q9LKW9	Q9LKW9 arabidopsis
43	254.5	9.2	575	10	Q9SLJ7	Q9SLJ7 arabidopsis
44	244	8.9	581	2	Q9HXX9	Q9HXX9 pseudomonas
45	241	8.7	581	2	Q9Z9H1	Q9Z9H1 pseudomonas

ALIGNMENTS

RESULT 1
ID Q9ZPK3 PRELIMINARY; PRT; 538 AA.
AC Q9ZPK3
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE SODIUM PROTON EXCHANGER NHX1 (FRAGMENT).
GN NHX1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=99145575; PubMed=9990049;
RA Gaxiola R.A., Rao R., Sherman A., Grisafi P., Alper S.L., Fink G.R.;
RT "The Arabidopsis thaliana proton transporters, AtNHX1 and AtPL1, can
function in cation detoxification in yeast."
RL Proc. Natl. Acad. Sci. U.S.A. 96:1480-1485(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Quintero F.J., Blatt M.R., Pardo J.M.;
RT "The AtNHX1 gene encodes a putative Na⁺/H⁺ antiporter."
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF106324; AAD16946.1; -.
DR EMBL; AF056190; AAF21755.1; -.
DR InterPro; IPR000676; -.
DR InterPro; IPR001179; -.
DR Pfam; PF00999; Na_H_Exchange; 1.
DR PRINTS; PR01084; NAHEXCHNGR.
DR PROSITE; PS00453; FKBP_PP1ASE_1; UNKNOWN_1.
FT NON_TER 538
SQ SEQUENCE 538 AA; 59513 MW; 1189AD6C5C726996 CRC64;

Query Match 100.0%; Score 2755; DB 10; Length 538;
Best Local Similarity 100.0%; Pred. No. 3.7e-169;
Matches 538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 VTILLISGKSHLLVSEDLFFIYLLPPIIFNAGFQVKKOFFRNFVTIMLFGAVGTII 120
 |||||
 Db 61 VTILLISGKSHLLVSEDLFFIYLLPPIIFNAGFQVKKOFFRNFVTIMLFGAVGTII 120
 |||||
 QY 121 SCTLISLGVTOFFKKLDIGTDFDGLDYLGAIFAATDSVCTLOVLNODETPLLYSLVFE 180
 |||||
 Db 121 SCTLISLGVTOFFKKLDIGTDFDGLDYLGAIFAATDSVCTLOVLNODETPLLYSLVFE 180
 |||||
 QY 181 GVVNDATSVVVFNAIQSFDTLHNLHAAAFHLLGNFLYLLFLLSTLLCAA----- 228
 |||||
 Db 181 GVVNDATSVVVFNAIQSFDTLHNLHAAAFHLLGNFLYLLFLLSTLLCAA----- 228
 |||||
 QY 229 TGLISAYVKKLYFG-----RHSTDEVALMMLAYLSYMLAELEFDSLGLTVFVFCG 280
 |||||
 Db 241 TGLISAYVKKLYFGPHINCHRHSTDEVALMMLAYLSYMLAELEFDSLGLTVFVFCG 300
 |||||
 QY 281 IVMSHYTHWNTYESSRITTKHTATLSFLAETFIYLVGMADLIDKWRVSVDTPGTSIA 340
 |||||
 Db 301 IVMSHYTHWNTYESSRITTKHTATLSFLAETFIYLVGMADLIDKWRVSVDTPGTSIA 360
 |||||
 QY 341 VSSILMGLVMVGRAAFVPLSFLSNLAKKNOSEKINFNMVQVIMWSGLMRGAVSMALAYN 400
 |||||
 Db 361 VSSILMGLVMVGRAAFVPLSFLSNLAKKNOSEKINFNMVQVIMWSGLMRGAVSMALAYN 420
 |||||
 QY 401 KTRAGHTDVRGNAMITSTIITVCLFSTVV 430
 |||||
 Db 421 KTRAGHTDVRGNAMITSTIITVCLFSTVV 450
 |||||

RESULT 4
 Q9FEB3 PRELIMINARY; PRT; 542 AA.
 AC Q9FEB3;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DE OSNHX1 PROTEIN.
 DE NA+/H+ EXCHANGER.
 GN PURPLE.
 OS Pharbitis nil (Violet) (Japanese morning glory).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
 OC Solanales; Convolvulaceae; Ipomoea.
 OX NCBI_TaxID=35883;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN-KK2K-2; TISSUE-LEAF, AND FULLY COLORED FLOWER BUDS;
 RX MEDLINE=20487008; PubMed=11034195;
 RA Fukada-Tanaka S., Inagaki Y., Yamauchi T., Saito N., Iida S.;
 RT "Colour-enhancing protein in blue petals.";
 RL Nature 407:581-581(2000).
 DR EMBL; AB033980; BAB16381.1; -
 DR EMBL; AB033989; BAB16380.1; -
 DR PRINTS; PRO1084; NAHEXCHNGR.
 SQ SEQUENCE 542 AA; 59973 MW; 4B47FDE04401A191 CRC64;

Query Match 74.1%; Score 2041; DB 10; Length 542;
 Best Local Similarity 73.8%; Pred. No. 2.3e-123;
 Matches 399; Conservative 57; Mismatches 71; Indels 14; Gaps 4;

QY 2 LDSLVSKLPSLSTSDHASVVALNLFVALLCACIVLGHLEENRWNESTALLIGLCTGV 61
 |||||
 Db 5 LSSLQN-SDLFTSDHASVSNRNLVALLCACIVLGHLEENRWNESTALLIGLCTGV 63
 |||||
 QY 62 TILLISGKSHLLVSEDLFFIYLLPPIIFNAGFQVKKOFFRNFVTIMLFGAVGTIIS 121
 |||||
 Db 64 VILLISGKSHLLVSEDLFFIYLLPPIIFNAGFQVKKOFFRNFVTIMLFGAVGTIIS 123
 |||||
 QY 122 CTTISLGVTOFFKKLDIGTDFDGLDYLGAIFAATDSVCTLOVLNODETPLLYSLVFE 181
 |||||
 Db 124 CTTISGAVKIFKHLIDIDFDGLDYLGAIFAATDSVCTLOVLNODETPLLYSLVFE 183
 |||||
 QY 182 VVNDATSVVVFNAIQSFDTLHNLHAAAFHLLGNFLYLLFLLSTLLGAATGLISAYVKKLY 241
 |||||

Db 184 VVNDATSVVVFNAIQSFDMTSDPKLIGLHIGNFLYLLFSLSTFELGVGIGLLCAVIKKLY 243
 |||||
 QY 242 FGRHSTDEVALMMLAYLSYMLAELEFDSLGLTVFVFCGIVMSHYTHWNTYESSRITTKH 301
 |||||
 Db 244 FGRHSTDEVALMMLAYLSYMAELFVLSGLTVFVFCGIVMSHYTHWNTYESSRITTKH 303
 |||||
 QY 302 TFATLSLAEFTFIYLVGMADLIDKWRVSVDTPGTSIAVSSILMGLVMVGRAAFVPLS 361
 |||||
 Db 304 SFATLSVAEFTFIYLVGMADLIDKWRVSVDTPGTSIAVSSILVGLVGRVAFVPLS 363
 |||||
 QY 362 FLSNLAKKNOSEKINFNMVQVIMWSGLMRGAVSMALAYNKFTAGHTDVRGNAMITSTI 421
 |||||
 Db 364 FLSNLAKKNOSEKINFNMVQVIMWSGLMRGAVSMALAYNKFTAGHTDVRGNAMITSTI 423
 |||||
 QY 422 TVCLFSTVVGMLTKPLSYLLP-----HONATTSMLSDDNTPKSIHPLLDQ--DS 471
 |||||
 Db 424 TVCLFSTVVGMLTKPLSYLLP-----HONATTSMLSDDNTPKSIHPLLDQ--DS 480
 |||||
 QY 472 FIEPSGNHNPVSDIRSGFTRTVHYVWRQFSDSPMRVFGVGRGFPVFGSPPTERN 531
 |||||
 Db 481 ESDMITGPEVARTALRMILLRTPTHTVHYWRKFDSPMRVFGVGRGFPVFGSPVQES 540
 |||||
 QY 532 P 532
 Db 541 P 541
 |||||

Query Match 73.4%; Score 2023.5; DB 10; Length 535;
 Best Local Similarity 73.3%; Pred. No. 3e-122;
 Matches 389; Conservative 60; Mismatches 79; Indels 3; Gaps 2;

QY 2 LDSLVSKLPSL-STSDHASVVALNLFVALLCACIVLGHLEENRWNESTALLIGLCTG 60
 |||||
 Db 3 MEVAAARIGALYTTSDVASVSNLNVALLCACIVLGHLEENRWNESTALLIGLCTG 62
 |||||
 QY 61 VTILLISGKSHLLVSEDLFFIYLLPPIIFNAGFQVKKOFFRNFVTIMLFGAVGTII 120
 |||||
 Db 63 VVILLMTKGSHLFFVPSDLFFIYLLPPIIFNAGFQVKKOFFRNFVTITLFGAVGTMI 122
 |||||
 QY 121 SCTLISLGVTOFFKKLDIGTDFDGLDYLGAIFAATDSVCTLOVLNODETPLLYSLVFE 180
 |||||
 Db 123 SFTTISNAIAFRSMNIGTLDVDFLGAIFAATDSVCTLOVLNODETPLLYSLVFE 182
 |||||
 QY 181 GVVNDATSVVVFNAIQSFDTLHNLHAAAFHLLGNFLYLLFLLSTLLGAATGLISAYIKKL 240
 |||||

```

Db 183 GYNDATSVLNFALONFDLHDAAVLVFLGNFFFLSLSTFLGVFAGLLSAYIIKKL 242
QY 241 YFGRHSTDEVALMMLMAYLSYMLAEFLDLSGILTVFPCGIVMSHYTHWNTVTESSRIITK 300
Db 243 YIGRHSSTDEVALMMLMAYLSYMLAEFLDLSGILTVFPCGIVMSHYTHWNTVTESSRIITK 302
QY 301 HPTATLSFLAEFIPLYVGMALDIDKRSVSDTPGTSIAVSSILMGLVMVGRAAFVPEL 360
Db 303 HATATLSFLAEFIPLYVGMALDIDKRSVSDTPGTSIAVSSILMGLVMVGRAAFVPEL 362
QY 361 SFLSNLAKKNOSEKINFNQVVMWGLMRGAVSMALAYNKFTFRAGHTDVRGNAMITST 420
Db 363 SFLSNLTKKAPNEKITWRQVVMWGLMRGAVSMALAYNKFTFRAGHTDVRGNAMITST 422
QY 421 ITVCLFSTVFGMLKPLISILYILP--HONATISM--LSDDNTPKSIHIPILLDDODSFTPEPS 476
Db 423 ITVCLFSTVFGMLKPLISILYILP--HONATISM--LSDDNTPKSIHIPILLDDODSFTPEPS 476
QY 481 VPRPDSIRGLTRPTVHYVWQFDDSFMRPVFGGRGVFPVPGSPTERN 531
Db 481 IVRPSSLRMLTKPTHTVHYVWQFDDSFMRPVFGGRGVFPVPGSPTERN 531

RESULT 6
Q9SQU0 PRELIMINARY; PRT; 552 AA.
AC Q9SQU0;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DE PUTATIVE SODIUM PROTON EXCHANGER.
GN P24P17.16.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
ON NCBI_TaxID=3702;
RX STRAIN=CV. COLUMBIA;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.,
RA Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome IIII BAC P24P17 genomic sequence.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC011623; AAF08577.1; -.
DR InterPro; IPR000676; -.
DR Pfam; PF00999; Na_H_Exchange; 1.
DR PRINTS; PR01084; NAHEXCHNGR.
DR PROSITE; PS00453; FKBP_LPTASE.1; UNKNOWN.1.
SQ SEQUENCE 552 AA; 61135 MW; FB4317D8A874FCE9 CRC64;

Query Match 69.3%; Score 1910.5; DB 10; Length 552;
Best Local Similarity 69.7%; Pred. No. 5.5e-115;
Matches 377; Conservative 58; Mismatches 97; Indels 9; Gaps 3;

QY 2 LDSLVKSLSTSDHASVVALNLFVALLCACIVLGHLLLEENRWNNESITALLIGLGTGV 61
Db 5 LSTMLEKTEALFASDHASVSVSNLNFVALLCACIVLGHLLLEENRWNNESITALLIGSCTGI 64
QY 62 TILLISKGKSSHLVSEDLFIYLLPIFNAGFQVKKQFFRNFVTMLFGAVGTIIS 121
Db 65 VILLISGGKSSRLVSEDLFIYLLPIFNAGFQVKKQFFRNFVTMLFGAVGTIIS 124
QY 122 CTIISLGVTOFQKLLDGTGLDYLATGAIFAATDSVCTQLVNLQDETPLYSLVFGEG 181
Db 125 FVITSGAKHLFEKNIGDLYLATGAIFAATDSVCTQLVNLQDETPLYSLVFGEG 184
QY 182 VYNDATSVVVEVNAIOSFDLTHLNEAFHLLGNFLYLLFLLSTLLGAAATGLISAYIVKIKLY 241

```

```

Db 185 VYNDATSVVLFNATQRELDNINSAIALEPAGNFFFLFLLSTALGVAAGLLSAFVKKLY 244
QY 242 FGRHSTDEVALMMLMAYLSYMLAEFLDLSGILTVFPCGIVMSHYTHWNTVTESSRIITK 301
Db 245 IGRHSTDEVALMMLMAYLSYMLAEFLDLSGILTVFPCGIVMSHYTHWNTVTESSRIITK 304
QY 302 TPATLSFLAEFIPLYVGMALDIDKRSVSDTPGTSIAVSSILMGLVMVGRAAFVPEL 361
Db 305 TPATLSFLAEFIPLYVGMALDIDKRSVSDTPGTSIAVSSILMGLVMVGRAAFVPEL 364
QY 362 FLSNLAKKNOSEKINFNQVVMWGLMRGAVSMALAYNKFTFRAGHTDVRGNAMITST 421
Db 365 FLSNLTKSPDEKIDLKQVTVWAGLMRGAVSMALAYNQFTTSQHTKVLGNAMITST 424
QY 422 TVCLFSTVFGMLKPLISILYILP--HONATISM--LSDDNTPKSIHIPILLDDODSFTPEPS 476
Db 425 TVCLFSTVFGMLKPLISILYILP--HONATISM--LSDDNTPKSIHIPILLDDODSFTPEPS 476
QY 477 GNNVPRPDSIRGLTRPTVHYVWQFDDSFMRPVFGGRGVFPVPGSPTERNPPDLS 536
Db 485 PRQHY---SFRMEFKSPSRAIHVYWRKFDNAYMRIRFGGRGVSPVFGSPVFGSPVQWS 540
QY 537 K 537
Db 541 E 541

RESULT 7
Q9FJ63 PRELIMINARY; PRT; 529 AA.
AC Q9FJ63;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DE SODIUM PROTON EXCHANGER.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
ON NCBI_TaxID=3702;
RX STRAIN=CV. COLUMBIA;
RA Nakamura Y., Sato S., Asamizu E., Kaneko T., Kotani H., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. VII.
RT Sequence features of the regions of 1,013,767 bp covered by sixteen
RT physically assigned P1 and TAC clones.";
RL DNA Res. 5:297-308(1998).
DR EMBL; AB015479; BAB08564.1; -.
SQ SEQUENCE 529 AA; 58871 MW; 8A6BD42E822F90C1 CRC64;

Query Match 56.9%; Score 1567; DB 10; Length 529;
Best Local Similarity 58.3%; Pred. No. 5.6e-93;
Matches 307; Conservative 80; Mismatches 120; Indels 20; Gaps 5;

QY 15 SDHASVVALNLFVALLCACIVLGHLLLEENRWNNESITALLIGLGTGVILLISKGKSSHL 74
Db 15 ACHPOVITSVFVAILCLVAILGHLLLEENRWNNESITALLIGLGTGVILLISKGKSSHL 74
QY 75 LVFSEDLFIYLLPIFNAGFQVKKQFFRNFVTMLFGAVGTIISLGVTOFQK 134
Db 75 LVFSEDLFIYLLPIFNAGFQVKKQFFRNFVTMLFGAVGTIISLGVTOFQK 134
QY 135 KLDTGTFDLGDLAIGATFATDSVCTQLVNLQDETPLYSLVFGEGVYNDATSVVFN 194
Db 135 KLDTGTFDLGDLAIGATFATDSVCTQLVNLQDETPLYSLVFGEGVYNDATSVVFN 194
QY 195 IQSPDLTHLNEAFHLLGNFLYLLFLLSTLLGAAATGLISAYIVKIKLYTCRHSSTOREVALM 254
Db 195 VQKIQFESLGTALQVFGNGLYLLFLLSTLLGAAATGLISAYIVKIKLYTCRHSSTOREVALM 254

```



```

QY 255 MLAYLSYMAELFDLSGLTVFVCGIWMVSHYTNHTVTESSRITTKHTFATLSPLAETFI 314
DB 255 VLMAYLSYMAELFDLSGLTVFVCGIWMVSHYTNHTVTESSRITSRVFAFLSPFAETFI 314
QY 315 FLYVGMALDIDKRSVSDTPGTSITAVSSILMGLVVMVGRAAAFVPLSFLSLAKKN--QS 372
DB 315 FLYVGMALDIDKRSVSDTPGTSITAVSSILMGLVVMVGRAAAFVPLSFLSLAKKN--QS 372
QY 315 FLYVGMALDIDKRSVSDTPGTSITAVSSILMGLVVMVGRAAAFVPLSFLSLAKKN--QS 372
DB 315 FLYVGMALDIDKRSVSDTPGTSITAVSSILMGLVVMVGRAAAFVPLSFLSLAKKN--QS 372
QY 373 EKINFMQVVIWWSGLMRGAVSMALAYKNKFTAGHTDVRGNAMITITVCLFSTVVFV 432
DB 373 EKINFMQVVIWWSGLMRGAVSMALAYKNKFTAGHTDVRGNAMITITVCLFSTVVFV 432
QY 375 ESITPKHQVVIWWSGLMRGAVSMALAYKNKFTAGHTDVRGNAMITITVCLFSTVVFV 434
DB 375 ESITPKHQVVIWWSGLMRGAVSMALAYKNKFTAGHTDVRGNAMITITVCLFSTVVFV 434
QY 433 MLTKPLISVLLP-----HQNAITSMUSDNTPK-SIHPLLDQDSEFSPGNHNVPRDPSI 487
DB 433 MLTKPLISVLLP-----HQNAITSMUSDNTPK-SIHPLLDQDSEFSPGNHNVPRDPSI 487
QY 435 FUTKPLVNLVLLPQDASHNTGNGRKTGPGSKEDATLPL-----SFDESASTNFNRKDSI 491
DB 435 FUTKPLVNLVLLPQDASHNTGNGRKTGPGSKEDATLPL-----SFDESASTNFNRKDSI 491
QY 488 RGLFTRPTTHVYWRQDDGFMRFVFGVGRGFPVPGSPPTERNPPD 534
DB 488 RGLFTRPTTHVYWRQDDGFMRFVFGVGRGFPVPGSPPTERNPPD 534
QY 492 SLLMEQPVYTIHRYWRKEDDYMRPIFG-----PRENOPE 528
DB 492 SLLMEQPVYTIHRYWRKEDDYMRPIFG-----PRENOPE 528

RESULT 8
Q9XZ44
ID Q9XZ44 PRELIMINARY: PRT; 561 AA.
AC Q9XZ44
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-NAR-2001 (TREMBLrel. 16, Last annotation update)
DE SODIUM-HYDROGEN EXCHANGER NHE1.
GN NHE1 OR CG12178.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Dow J.A.T.;
RT "An insect member of the Na+/H+ exchanger family";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF142676; AAD32689.1;
DR FlyBase: FBgn026787; Nhe1.
DR InterPro: IPR000676;
DR Pfam: PF00999; Na_H_Exchange; 1.
DR PRINTS: PR01084; NAHEXCHNGR.
DR SEQUENCE 561 AA; 60974 MW; 6B53FC7A92528627 CRC64;

Query Match 22.3%; Score 615.5; DB 5; Length 561;
Best Local Similarity 33.6%; Pred. No. 6.4e-32;
Matches 163; Conservative 93; Mismatches 148; Indels 81; Gaps 18;

QY 10 PSLTSDHASV-----ALNLFVALLCACIVLGHLEEN 43
DB 63 PSISASGNASTTKRGNASTLVTPDPLDSDHVEGHNHNSLSLF-VTCV-IMLGILLHS 120
QY 44 -----RWNNESITALLIGLGTGVITLLIS--KGKSSHLVFESEDLFFYLLPPIIFNAG 95
DB 121 MLQTGFQYLPESIVVFLGAFGLSLNVMSGONGSKRKEEVFSPMGFLVLLPPIIFESG 180
QY 96 FOVKKKQFPRNVVTMLGAVGTIISCTIISGLVTPQFKLDIG-----TDLGDYLAIGA 151
DB 181 YNLHKGNFQNTGSLVFAIFGTTISALVIGAGI-----YLLGLGEVAFRLSFSSESFAFGS 236
QY 152 IFAATDSVCTLOVNO-DTEPLLYSLVFGVGVNDATSVVFNVAI-OSFDLTHLNHEAA- 208
DB 237 LISADVPATVAIFHALVDPLNMLVFGESILNDATISIVTASITOS---ANNVNAEAST 293
QY 209 ----PHLLGNFLYLLLSLLGAATGLISAYVKKLYFGFRHSTREVALMLMAYLSYML 264
DB 294 GEAMFSALKTCFCAEFASAGIGVIFALISALLKHLIDLRKHP-SLEFAMLMFTYAPYL 352

```

```

QY 265 AELFDLSGLTVFVCGIWMVSHYTNHTVTESSRITTKHTFATLSPLAETFI 324
DB 353 ABGILGIMAILFCGIVMSHYTNHTVTESSRITTKHTFATLSPLAETFI 412
QY 325 IDKWSVSDTPGTSITAVSSILMGLV--MVGRAAFVPLSFLSLAKKNQSEKINENMOVV 382
DB 413 FKH-----QVELSFVIAVLCIGRACNIFPLAFLVN---KFRCHKNNKMOFI 459
QY 383 IWSGLMRGAVSMALAYKNKFTAGHTDVRGNAMITITVCLFSTVVFVGLTKPLI 439
DB 460 MWFSGL-RGAIYSALSL-----HLNLDSEKRRHVITITLIIVLTFLVGLSGSTPMLL 511
QY 440 SYLLP 444
DB 512 KYLKP 516

RESULT 9
Q9VPJ1
ID Q9VPJ1 PRELIMINARY: PRT; 649 AA.
AC Q9VPJ1;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE NHE1 PROTEIN
GN NHE1 OR CG12178.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Balieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy C., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pattman G.S., Pan S., Poillard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamis I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).

```

Query Match 21.3%; Score 585; DB 10; Length 140;
Best Local Similarity 80.0%; Pred. No. 1e-30;

100 KKQFERNFTIMLFGAVGTIISCTIIISLGVTQFFKKLDIGTDFDLGDIAGIAFAATDSV 159
|||||:::|||||::|||::|||::|||::|||::|||::|||::|||::|||
1 KKQFERNFITIVFGALGVLSFTTIISLGAEFFFKLDIGSLEGLDYLAIAGAFAATDSV 60

160 C T L Q V I N Q D E T P L Y S I V F G E G V N D A T S V V F N A I Q S F D L T H L N H E A F H L I G N F L Y L F 219
 | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

61 C T L O V L H O D E T P L Y S I V F G E G V N D A T S V L F N A I Q N E D T H I D R I A K F A G N F L Y L F 120

220 LLSTLLGAATGLISAYVIKK 239
 ||||| |||:|:|:|
 121 FTSTLLGAATGLISAYVIKK 140

RESULT 11
XW14

Q9AW14 FREELIMINARI, FBI, 703 AA.
Q9XW14;
01-NOV-1999 (TReMBLrel. 12, Created)
01-NOV-1999 (TReMBLrel. 12, Last sequence update)

Y18D10A.6 PROTEIN.
Y18D10A.6.
Y18D10A.6.
Caenorhabditis elegans.
Caenorhabditis elegans.

Eukaryota; Metazoa; Nematozoa; Chironomidae; Nematoda; Rhabditiidae; Peloderinae; Caenorhabditis.
NCBI_TaxID=6239;
[1]

SEQUENCE FROM N.A.
A Harris B.;
Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
[2]

SEQUENCE FROM N1A.
MEDLINE=94150718; PubMed=7906398;
Willson R., Ainscough R., Anderson K., Baynes C., Berks M.,
Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,

Craxton M., Dear S., Du Z., Durbin K., Favellio A., Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L., Jones M., Kershaw J., Kirsten J., Laister N., Latreille P., Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,

Parsons J., Percy C., Kliesen L., Koopra A., Saunders D., Snowkneel K., Smailson N., Smith A., Sonhammer E., Staden R., Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.

2.2 kb of contiguous nucleotide sequence from chromosome III of *C. elegans*.²; Nature 368:32-38(1994).
EMBL: AF034393; CAA22320.1; -

InterPro: IPR0000676; -
 Pfam: PF00999; Na_H_Exchange; 1.
 PRINTS: PR01084; NAHEXCHNG.
 SEQUENCE 703 AA: 76942 MW: 9054503CERED767 CR664;

Query Match	21.28;	Score 583.5;	DB 5;	Length 703;
Best Local Similarity	34.6%;	Pred No	9.3e-30;	

Matches 150; **Conservative** 79; **Mismatches** 160; **Indels** 45; **Gaps**

25 LFVALLCACIVLGHLLLEENRWMNESITALLIGTGTGVTILLISKGRSSHLVFSEDLFFI 84
||| . | . . . |||. . . | . || | : : : | : | : | : | : | :

138 LFVIMLATLVHMLIVSKLHHMPESDAIVALGALIG-SILSYSRKDWSEIEALSPDVFFL 190
85 YLLPPIIFNAGFQVKKKQFFERNFVTIMLFGAVGTIISCTIISLGYTQFFKKLDIGTFDLG 144

19/ VLLPPLIFENAYNLNKGYFFSNFVPLLTFALEGTTSANVIGAGLYILGAIGLFEFTTF 250

1

"The genome sequence of Drosophila melanogaster.";

RT Science 287:2185-2195(2000).
RL EMBL; AE003614; AAF52423.2; -.
DR FlyBase; FBgn0028703; Nhe3.
DR InterPro; IPR000676; -.
DR Pfam; PF00999; Na_H_Exchange; 1.
DR PRINTS; PR01084; NAEXCHNGR.
SQ SEQUENCE 727 AA; 80331 MW; 36D730CC7C3B17B CRC64;

Query Match 20.7%; Score 570; DB 5; Length 727;
Best Local Similarity 29.3%; Pred. No. 7.1e-29;
Matches 178; Conservative 109; Mismatches 189; Indels 132; Gaps 24;

QY 9 LPSLSTSDHASVVALNL--FVALLCACIVLGHLLLENR--WNESITALLIGL----- 57
Db 65 LQAKATLNIH-RIQSLDLLVFVLLALTTLTWLFKHHRSWMLHETGLAVTYGLVIGAIIR 123
QY 58 --GTGVTTILLIS-----KGRSSHLLV----- 76
Db 124 YAGTSATLHVHMQVEPQGVPTYSKLPDPDILWFRYPVNTQNGTKPPEGIKTYAYVFRGOVH 183
QY 77 -----FSEDLFFIYLLPIIFNAGFOVKKKOFFRNFVITMLFGAVGTIISCTI 124
Db 184 DVDENEIDLKATDPPEVFFNIIPLPIIFVAGYSLKKYFFERNLGAITFAIVGTTLSAFL 243
QY 125 IS---LGVTOFFKLDIGTDFLDGDLAIGAFAAATDSVCTLOVLNODETPL-LYSLVFGE 180
Db 244 IGGFMVGCVKLMPKYLSSSFTFLDSLYFGALISPTDPLTLAIFNDLRVDVNLVALVLGE 303
QY 181 GVNDAATSVVFNAIQSFDTLHNLHAAFHL-----LGNFLYLLFLLSLGAAATGLISA 234
Db 304 SVLNDVAIVLSGAIQNYG-EHYSNTGEFETAPLRSLSDFFSIFLLSLMIGAAMGLTA 362
QY 235 YVIKKLYFGRHSTDREVALMMLMAYLSYMLAEFLDLSGLITVFFCGIVMSHYTHWNVTES 294
Db 363 -LMTKTRVRDPFLLSALFVLMSTYFLAEATELTGVVAVLFCGICQAHYTHNNLSGD 421
QY 295 SRITTKHTFATLSLAETFFIFYVGMALDIDKWRVSVDTPGTSIAVSSILMGLV--MVG 352
Db 422 SRQTKQIFELLNLAENFIYSIGV-----SMFTPKHHFDAGFIITAFICAAIG 472
QY 353 RAAFVFLPSLNLAKNOKSEKINFNQVIVWWSGLMRGAVSMALAYNKFTTRAGHTDVRG 412
Db 473 RAVNVYPLSWLLNKKR---PKISTNFQHMLFFAGL-RGAMSFALA-----IRNTVSDER- 523
QY 413 NAIMTITIVCLFSTVVFGLTKPLISYL-LPHONATTSMLSDDDNTPKSIHIPLLDQDS 471
Db 524 -QTMLTATSLIVFTVVIQGAANFLNLWKIP-----VGVDDTEQLNNYQVHSSDG 575
QY 472 FIEP-----SG--NHNVPRP-DSIRGFLTRPT-----RTVH-----YYWRQF 505
Db 576 YLQDVEGGVGNRNKRLSGGTDNLDTVPDGTNGSLGGASGRRRNSHEKAILARIWGNF 635
QY 506 DDSFMRPV 513
Db 636 DTKVMKPL 643

RESULT 14
Q9UG24 PRELIMINARY; PRT; 687 AA.
AC Q9UG24;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE SODIUM-HYDROGEN EXCHANGER NHE3.
GN NHE3 OR CC11328.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_taxid=7227;

[1]
RN SEQUENCE FROM N.A.
RP Giannakou M.E., Dow J.A.T.;
RT "A novel Drosophila member of the NHE exchanger family.";
RL Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF199463; AAF13702.1; -.
DR FlyBase; FBgn0028703; Nhe3.
DR InterPro; IPR000676; -.
DR Pfam; PF00999; Na_H_Exchange; 1.
DR PRINTS; PR01084; NAEXCHNGR.
SQ SEQUENCE 687 AA; 75775 MW; C5D6BA9BDDCBA1DE CRC64;

Query Match 20.6%; Score 568; DB 5; Length 687;
Best Local Similarity 29.3%; Pred. No. 8.9e-29;
Matches 178; Conservative 105; Mismatches 193; Indels 132; Gaps 23;

QY 9 LPSLSTSDHASVVALNL--FVALLCACIVLGHLLLENR--WNESITALLIGLGTGVTIL 64
Db 25 LQAKATLNIH-RIQSLDLLVFVLLALTTLTWLFKHHRSWMLHETGLAVTYGLVIGAIIR 83
QY 65 LI-SKGRSSHLLV----- 76
Db 84 YAGTSGLDHHMQVEPQGVPTYSKLPDPDILWFRYPVNTQNGTKPPEGIKTYAYVFRGOVH 143
QY 77 -----FSEDLFFIYLLPIIFNAGFOVKKKOFFRNFVITMLFGAVGTIISCTI 124
Db 144 DVDENEIDLKATDPPEVFFNIIPLPIIFVAGYSLKKYFFERNLGAITFAIVGTTLSAFL 203
QY 125 IS---LGVTOFFKLDIGTDFLDGDLAIGAFAAATDSVCTLOVLNODETPL-LYSLVFGE 180
Db 204 IGGFMVGCVKLMPKYLSSSFTFLDSLYFGALISPTDPLTLAIFNDLRVDVNLVALVLGE 263
QY 181 GVNDAATSVVFNAIQSFDTLHNLHAAFHL-----LGNFLYLLFLLSLGAAATGLISA 234
Db 264 SVLNDVAIVLSGAIQNYG-EHYSNTGEFETAPLRSLSDFFSIFLLSLMIGAAMGLTA 322
QY 235 YVIKKLYFGRHSTDREVALMMLMAYLSYMLAEFLDLSGLITVFFCGIVMSHYTHWNVTES 294
Db 323 -LMTKTRVRDPFLLSALFVLMSTYFLAEATELTGVVAVLFCGICQAHYTHNNLSGD 381
QY 295 SRITTKHTFATLSLAETFFIFYVGMALDIDKWRVSVDTPGTSIAVSSILMGLV--MVG 352
Db 382 SRQTKQIFELLNLAENFIYSIGV-----SMFTPKHHFDAGFIITAFICAAIG 432
QY 353 RAAFVFLPSLNLAKNOKSEKINFNQVIVWWSGLMRGAVSMALAYNKFTTRAGHTDVRG 412
Db 433 RAVNVYPLSWLLNKKR---PKISTNFQHMLFFAGL-RGAMSFALA-----IRNTVSDER- 483
QY 413 NAIMTITIVCLFSTVVFGLTKPLISYL-LPHONATTSMLSDDDNTPKSIHIPLLDQDS 471
Db 484 -QTMLTATSLIVFTVVIQGAANFLNLWKIP-----VGVDDTEQLNNYQVHSSDG 535
QY 472 FIEP-----SG--NHNVPRP-DSIRGFLTRPT-----RTVH-----YYWRQF 505
Db 536 YLQDVEGGVGNRNKRLSGGTDNLDTVPDGTNGSLGGASGRRRNSHEKAILARIWGNF 595
QY 506 DDSFMRPV 513
Db 596 DTKVMKPL 603

RESULT 15
O13726 PRELIMINARY; PRT; 569 AA.
AC O13726;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE PUTATIVE SODIUM/HYDROGEN EXCHANGER C15A10.06.
GN SPAC15A10.06.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_taxid=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Murphy L., Harris D., Wood V., Barrell B.G., Rajandream M.A.;
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -1- SIMILARITY: WITH OTHER NA(+)/H(+) EXCHANGER ISOFORMS, PARTICULARLY
 CC IN THEIR N-TERMINAL HALVES.
 DR EMBL: Z97208; CAB10103.1; -.
 DR InterPro: IPR000676; -.
 DR Pfam: PF00999; Na_H_Exchange; 1.
 DR PRINTS: PR01084; NAHEXCHNGR.
 KW Hypothetical protein; Transmembrane; Glycoprotein; Sodium transport;
 KW Transport; Symport.
 FT TRANSMEM 34 54 POTENTIAL.
 FT TRANSMEM 67 87 POTENTIAL.
 FT TRANSMEM 90 110 POTENTIAL.
 FT TRANSMEM 131 151 POTENTIAL.
 FT TRANSMEM 162 182 POTENTIAL.
 FT TRANSMEM 191 211 POTENTIAL.
 FT TRANSMEM 230 250 POTENTIAL.
 FT TRANSMEM 280 300 POTENTIAL.
 FT TRANSMEM 339 359 POTENTIAL.
 FT TRANSMEM 395 415 POTENTIAL.
 FT TRANSMEM 428 448 POTENTIAL.
 FT DOMAIN 15 22 POLY-ASP.
 FT DOMAIN 428 431 POLY-VAL.
 SQ SEQUENCE 569 AA; 63911 MW; 3912BC6CBF68FE4D CRC64;

Query Match 19.7%; Score 544; DB 3; Length 569;
 Best Local Similarity 31.7%; Pred. No. 2.5e-27;
 -Matches 138; Conservative 103; Mismatches 157; Indels 38; Gaps 15;

QY 22 ALNLFVALLCACIVLGHLEEN--RMNESITALLIGLGTGVTILLISKGR--SSHLVFS 78
 DB 36 ALFILLVLLGALLTSYVQSKKIRAIHETVTSVFGVWVGL-IIRVSPGLIIONMVSFH 94
 QY 79 EDLFYLLPPIIFNAGFQVKKQFRNPVTIMFGAVGTIISCTIISLGVTF----- 132
 DB 95 STYFFNVLLPPIILNSGYELHOSNFRNIGTITLTFAGTFISA--VTGLVLYVIFSF 152
 QY 133 FKLDIGTFDGLDYLAIGAIFAATDSVCTQVLNDET-PLIYSLVFGGVNDATSVV 191
 DB 153 FENLSM-TF--VEALSMGATLSATDPVTVAIFNSYKVDQKLYTIIFGESILNDAVIM 209
 QY 192 FNAIQSFDLTHLNEAAPHLLGNFLYFLLLGAAATGLISAYVKKLYFGRHSTDREV 251
 DB 210 FETLQOFQGGKTLHFTFLFSGIGFIITFFISLLIGVSGITALLKYSYLRYPY-IES 268
 QY 252 ALMMLMAYLSYMLAEFLDLSGILTVFCGIVMSHYTHWNTVTESSRITTKHTFATLSFLAE 311
 DB 269 CIULLMAYTSYFFSNGCHSGVSVLLFCGILKHYAFFNMSYKAKLSTKYVFRVLAQLSE 328
 QY 312 TFIFYVGMAL--DIQKRSVSDTPTSTAVSSILMGLVMVGRAAFVPLSFLSLAKK- 369
 DB 329 NFIFYLGSLSFTQVDLVYKPIFILITTVAVTA-----SRVMNVEPLSNLLNKHFRQ 380
 QY 370 ---NQSEKINFNQVVIWNSGLMRGAVSMALAYNKFTAGHTDVRGNAMITSTITVCLF 426
 DB 381 RGNLIDHIPYSYQOMLFWAGL-RGAVGVALA-----AGEGENAQTLRAT-TLVVVVL 432
 QY 427 STVVFGLTKPLISYL 442
 DB 433 TLIFGTTARMLEIL 448

RESULT	6
TM021B04/c	
LOCUS	
DEFINITION	TM021B04_90019 bp DNA
ACCESSION	Arabidopsis thaliana BAC TM021B04.
	AF007271
	PLN
	12-JUN-1997

VERSION AF007271.1 GI:2191181

KEYWORDS

ORGANISM

thale cress.
Arabidopsis thaliana
Eukaryota: Viridiplantae: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis.

REFERENCE

1 (bases 1 to 90019)
Dante, M., Wamsley, P. and Gibson, A.

JOURNAL
The sequence of A. thaliana TM021B04
Unpublished (1997)

REFERENCE
2 (bases 1 to 90019)
Washington University Genome Sequencing Center.

JOURNAL
The A. thaliana Genome Sequencing Project
Unpublished (1997)

REFERENCE
3 (bases 1 to 90019)
Waterston, R.

Submitted (06-JUN-1997)
Submitted by:

Genome Sequencing Center
Department of Genetics, Washington University,
St. Louis, MO 63108, USA
e-mail: twilson@watson.wustl.edu

MAPING: Clones were assigned to the YAC map by hybridization by M. Iohdi, Cold Spring Harbor Laboratories, and fingerprinted by M. Marra, WashU, to pick the best candidates for sequencing.

NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone

NEIGHBORING COSMID INFORMATION:

The actual start of this clone is at base position 1 of TM021B04 ; the actual end is at base position 90019 of TM021B04. The orientation of this clone is unknown.

NOTES:

Coding sequences below are predicted from computer analysis, using the program Genefinder (P. Green and L. Hillier, ms in preparation). Location/Qualifiers

FEATURES
source
1. 90019
/organism="Arabidopsis thaliana"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="V"
/clone="TM021B04"
/map="5"
1802..3685
/gene="A_TM021B04.11"
join(1802..2050,2468..2686,2758..2844,2901..3095,3375..3685)
/gene="A_TM021B04.11"
/note="contains similarity to synecococcus PCC7942 chromosomal region used as basis of neutral siteit recombination cloning vector (PID:g1174192)"
/codon_start=1
/evidence="not_experimental"
/protein_id="AAB61082.1"

gene

CDS

gene

CDS

gene

CDS

/db_xref="GI:2191197"
/translation="MWMILHSVGLIPCSNQOKSFHSHYRRCIVCSSTFGSLIRRO
ALEVDYSKLSGGERALSLVQDLGKPGGLCFGAARQVQRLYTEELKLGINAA
SLSPPTTGSGIEREQLIAVSGGIVAKRAEDLSQOFLTLGLGFWLTLVGFN
GISTGIVLDTGTGTFEESORHNRVYOGKHVITFMWYCELRSGHEGHLVAVLGLR
GYTSLISLEALQEGSLINIQAGSAFVDEYLEPKNKKICLLPQNMNLVRSFCILAAVA
PBYLLXGABEGGLDDISKSFELPLKNSDYNNHLYGVYVIMEDLDVYSLGFLPQK
ADQVRMSVNLITLLRHEIARSKLAQMSGSEVSGCIOTIDSDIDPDI
complement(3906..4727)
/gene="A_TM021B04.14"
complement(join(3906..4337,4418..4497,4601..4727))
/gene="A_TM021B04.14"
/note="coded for by A. thaliana cDNA T22079"
/codon_start=1
/protein_id="AAB61076.1"
/db_xref="GI:2191191"
/translation="MNAFSSNAATISSSSSPLSIPSPKRRQSSPPRIYLSNKR
EDKIDPDQHSSESSSLFRRTISNDENGLYLSAASKYFTTSCHEGSLPARDY
DVSTFTWSLFTKSPRRMRVAPTCVCGQRTIRLINFHAYTDFVYVQCGGNFHK
LVNINLFEHVKITVSSSSSFDYDAKVDVSGILFDHDDDDNADSDVPEPL"
complement(12650..16625)
/gene="A_TM021B04.10"
complement(join(12650..13129,13415..13509,14187..14498,
14596..16014,16127..16625))
/gene="A_TM021B04.10"
/note="similar to N. tabacum membrane-associated
salt-inducible protein (PID:g473874)"
/codon_start=1
/evidence="not_experimental"
/protein_id="AAB61077.1"
/db_xref="GI:2191192"
/translation="MKSDFTSTYTHNPSTFLPKIPSRNSRISIKSSSSSKVRDDPW
SLSDNDEPKPRYERKPHLSDDDAIRLIRKKAQYLSTLRNQGSAAMPKIRTP
EQMVOYLEDNDQMGYKRVAAIKVRLRSLQRSGSDMFMSVFAKSLSPFVACV
VLKRGFCVYKOLSPRSVVVYIVLRIYGVGKIKMAEFLEMLEVCGEPDAVAG
TLCCTYAKRGHSAMLFFYKAVOERILISTSVYFPMSSLOKSPFGKVIDMLMEV
EEGAPNPEYTVTVSSYAKCGFERALAPFEMSLGFEVETTSVYSLSVAGD
WEKAGTEEMRSGCIYPSWYTCNPMILSTYVETPEPALSIFPMERKIPDQVIR
GLIRIYKGLDFHDAQSMFEYTERNLADKITYLANSQYHLSNGVYKALDYEM
KIRDPISRRITVYMLQCTAKTQVDCAEARLKEGSAVADQLTVNKGREARVADNFY
GKAAGFKQIOWDVQVDFELKTMARYKREKSGVAAQDQLTVNKGREARVADNFY
QTLSAENITVNRHDEAVLVANSQIDLVNMLGLMLRLKEGNLNEKAILNLEKTDL
GSSAVNRYISFEVREGDVSRAEMLADIIIRGLRMEETIATLLAVGRQHKLEAKR
LYLAAGESKPKGKSVIRSMIDAYVRCGMLEDAVGFMSAGAKICATSRHLHVEDLL
TKRGLIDEEKIYTNMIMHMGGMYLKIKFIMMLRPSYNMAVKICATSRHLHVEDLL
QAMERNRCRDLSTYFLTIQVYAESQFPAEKTYTLVKEKIDPLSHSHSSLSLAV
KAGMERAEARTYCKMSSEAGISPDASCKRTILKGYTCGDARKGILPEKMIRESVEDD
PVSQVYVDLYKAKGKRDY"
complement(17345..24723)
/gene="A_TM021B04.9"
complement(join(17345..17911,17988..18340,19797..20049,
20117..20610,21561..24723))
/gene="A_TM021B04.9"
/note="contains similarity to a DNMJ-like domain"
/codon_start=1
/evidence="not_experimental"
/protein_id="AAB61072.1"
/db_xref="GI:2191187"
/translation="MDMKNKPEARAKTLEADKMKRGDPVGAOKILLKAOSSLFSGEST
RMAAYGDVHNSAEKTKNCKLEAMYGILQYMHKRAADAPITLALALLIPKDNQST
AEARKLVMDASRFLADKRSQYDIBRIYIYRLTNLNNSSLOCAATNSITDFW
TCEBNGSIRIKYLRKYVAILNLCNITQSIYMAVDFEAPASKSTQKSVQMGPCN
TSLNNGESITGAQGSVAEVDRKGFNNKPRKNGSGESKRTYVKNKSLTEFVDF
DEKKMSAKLHKPHPEVTEPELGASQVSDGSVRSIDNSLIGPKPKRSLNGLVESI
INDGSDFLKDAVSKDNKSKSPRRSQGSVSEBEKIDNSLIGPKPKRSLNGLVESI
OYTRKGLVAGSSKRKEDSGGSAPCAKRAKFFVSGYQOELISVDEKERRAS
EGGVYAEKIDNNRPNENITFDLPDPFSEFETLTCFQGNQWMAVYDIPGMPL
YARIDYVPEPKLMTWIDPDQDNKNSIPLAGIFQGGGSEENDHLPSCOMPL
TRNSVYVYVPEKGIATIPRGDLSWSASENHKRPYDVEVLELSPENDNLSIGF
LAKYDCEFSLFRDADQGVQIDIPRQSLRSHVSPFKTKGREGVPGCFEDP
ALPELEFVYDVKVYGVGDRELDNPKGTGKPPPEKSKVPRLLSTNSQANSRGR
HDNRSFESKESDGRINHERNSAKESRSKRVKVDLNLNRKSPRLSTNSQANSRGR
QIAEKSSANNESCGKPGGLISGEMSEKITTRRKRTPODIDRPLVNVKRGNDSE

